

# Energy Harvesting in Secure MIMO Systems

Wei Wu, *Member, IEEE*, Xueqi Zhang, Baoyun Wang, *Member, IEEE*,

**Abstract**—The problems of energy harvesting in wireless communication systems have recently drawn much attention. In this paper, we focus on the investigation of energy harvesting maximization (EHM) in the important secrecy multi-input multi-output (MIMO) systems where little research has been done due to their complexity. Particularly, this paper studies the resource allocation strategies in MIMO wiretap channels, wherein we attempt to maximize the harvested energy by one or multiple multi-antenna energy receivers (ERs) (potential eavesdropper) while guaranteeing the secure communication for the multi-antenna information receiver (IR). Two types of IR, with and without the capability to cancel the interference from energy signals, are taken into account. In the scenario of single energy receiver (ER), we consider the joint design of the transmission information and energy covariances for EHM. Both of the optimization problems for the two types of IR are non-convex, and appear to be difficult. To circumvent them, the combination of first order Taylor approximation and sequential convex optimization approach is proposed. Then, we extend our attention to the scenario with multiple ERs, where the artificial noise (AN) aided weighted sum-energy harvesting maximization (WS-EHM) problem is considered. Other than the approaches adopted in solving the EHM problems, an algorithm conducts in an alternating fashion is proposed to handle this problem. In particular, we first perform a judicious transformation of the WS-EHM problem. Then, a block Gauss-Seidel (GS) algorithm based on logarithmic barrier method and gradient projection (GP) is derived to obtain the optimal solution of the reformulation by solving convex problems alternately. Furthermore, the resulting block GS method is proven to converge to a Karush-Kuhn-Tucker (KKT) point of the original WS-EHM problem. Simulation results show the outstanding energy harvesting performance and computational efficiency of our proposed algorithms.

**Index Terms**—Energy harvesting, physical-layer security, resource allocation, MIMO wiretap channels.

## I. INTRODUCTION

WITH the rapid development of wireless communication technology and the incremental user requirement, the ubiquitous high data rate becomes a basic merit of our next generation wireless communication networks, especially in the future 5G system. However, it indeed implicits a rapidly escalated energy consumption of the communication nodes serving in their corresponding networks to maintain such a property. Consequently, the green communication technologies for the purpose of decreasing the energy dissipated by

each unit bit information transmission as much as possible attract much attention among researchers [1]–[9]. Although the energy efficiency appears to be enhanced significantly by bringing in these advanced technologies, the limited lifetime of the wireless networks still remains as a bottleneck of its own development since the mobile devices are powered by batteries with limited action time. In general, one may recharge or replace the batteries to prolong the lifetime of the mobile devices. Nevertheless, it would lead to not cost-effective, inconvenient (say, in harsh environments or transducers embedded in human bodies or buildings) or environmentally-unfriendly. Hence, a cost-efficient, convenient, as well as eco-friendly alternative which can assist in capturing energy from surroundings is proposed. Other than the traditional sources such as solar, water conservancy, and biomass, etc., the ambient radio frequency (RF) signals turn out to be a feasible new source used to scavenge energy.

Obviously, it is a promising alternative to incorporate energy harvesting into wireless communication devices, which not only provides the perpetual renewable energy but also it is particularly suitable for low-power communication nodes. Many researchers from all over the world paid much attention to the RF-based energy scavenging systems in recent decades. The original recorded study of this field can be traced back to the 1914s in Nikola Tesla's work [1], where some preliminary attempts of wireless power transfer about sensor networks, body area networks and passive RF identification systems are demonstrated. On the other hand, in accordance with recent research literatures [3]–[9], the RF-based energy transfer technology can facilitate simultaneous wireless information and power transfer (SWIPT) since the signals used to carry energy are able to serve as a vehicle for message transportation. Such a discovery brings numerous new challenges and opportunities to the engineers in this field. Authors in [3] first came up with a capacity-energy function and studied the essential tradeoffs between information decoding rate and power transfer for the single-input single-output (SISO) systems. Literature [4] extended the situation in [3] to a frequency selection single antenna additive white Gaussian noise (AWGN) channels, and proposed an efficient resource allocation algorithm for the near-field communication systems they concerned. The works in [5]–[7] further pursued the multi-input single-output (MISO) systems with multi-antenna transmitters. By exploiting the available degrees of freedom offered by multi-antenna transmitter, the energy transfer efficiency can be improved distinctly. The common three nodes MIMO broadcast channels were investigated in [8] and [9]. As a compromise, Zhang et al. [8] presented two practical schemes, i.e., time switching and power splitting, for SWIPT. In [9], a transmission covariance optimization strategy was studied for the EHM problem.

From the papers outlined above, a consensus was revealed

Wei Wu is with the College of Communication and Information, Nanjing University of Posts and Telecommunications, Nanjing, 210003, China (e-mail: bywunjupt@foxmail.com).

X. Zhang and B. Wang are with the College of Communication and Information, Nanjing University of Posts and Telecommunications, Nanjing, 210003, China (e-mail: zhangxueqi9231@163.com, bywang@njupt.edu.cn).

This paper was supported by the National Natural Science Foundation of China (No. 61271232, 61372126, 61501171); the Open research fund of National Mobile Communications Research Laboratory, Southeast University (No. 2012D05); the Priority Academic Program Development of Jiangsu Province (Smart Grid and Control Technology) for funding.

that more transmit power will be consumed for the purpose of higher energy scavenging at ER. However, in fact, an increasing transmission power means a higher susceptibility to eavesdropping due to the result of a substantial information leakage and the broadcasting nature of channels. Therefore, how communication security can be ensured arises as a crucial issue in SWIPT systems.

Recently, physical-layer security draws much attention in wireless communications. Different from the traditional way employing cryptographic encryption at the application layer, physical-layer security provides perfectly security based on the knowledge of physical properties of wireless channels. In the 1970s, Wyner who first studied physical-layer security and proved a positive secrecy rate for degrade channels in [10]. Notable extension of Wyner's work included in [11] and [12], where the non-degraded channels and the SISO Gaussian wiretap channels are considered, respectively. Furthermore, from the perspective of information theory, Loyka et al. [13] obtained the optimal transmit covariance matrix over general three-node MIMO secrecy channels (not just the special channels mentioned in [11] and [12]) and gave a direct proof of necessary condition of optimality based on the necessary Karush-Kuhn-Tucker (KKT) conditions. A custom-designed algorithm taking advantage of extended barrier method was proposed in [14]. Li et al. [15] designed an AN-aided resource allocation strategy to maximize the achievable secrecy rate for an MISO channel since AN has been found to be a promising tool for preventing eavesdropping. Then, [16] extended their works to the MIMO case, where an alternating optimization (AO) method to handle the secrecy optimization problems was derived through a judicious SCM (secrecy capacity maximization) reformulation. Wang et al. [17] considered a two-hop AF relay network, a generalized singular value decomposition (SVD) based precoding approach was proposed to acquire the optimal power allocation for secrecy capacity. Besides, it is also note that the secure communication in SWIPT has been widely studied in the literatures [18]–[25], where the trade-off designs between secrecy rate and energy scavenging are achieved in different situations.

However, it is worth to note that the aforementioned works focusing on energy harvesting while meeting the secure communication requirement at IR are limited to the MISO systems or even the SISO systems. The more complicate MIMO systems still remain less studied in this field (most are confined to secrecy rate or energy efficiency maximization), though a preliminary work has been done in [9]. In this paper, we consider the energy harvesting problems for MIMO secrecy systems with one or multiple multi-antenna ERs. Our problem formulations contain two types of IR, namely, the IR with and without capability to perform energy signals interference cancellation. The main contributions of this paper are summarized as follows.

- For the one ER case (a common three node communication system), two EHM problems with the optimization of transmit covariance matrixes are considered. The formulated optimization problems, denoted by (P1) and (P2), are non-convex and subject to secrecy rate constraint for IR and given total power budget of the transmitter. By

resorting to the Taylor series expansion, we reformulate both problems into convex ones. After that, for (P1), we further convert it into a two-layer optimization problem. Then, the iterative WF-SVD (water filling and SVD) algorithm resulting in the semi-closed form solutions and the golden section search algorithm are proposed for the inner layer and outer layer optimizations, resp.. For (P2), the off-the-shelf convex optimization tool, such as SeDuMi [26] (also called as CVX [27]), is utilized to obtain its optimal solution.

- For the multiple ERs case, the AN-aided WS-EHM problem is considered. And, different from above case, since the distinction caused by two types IR turn out to be obscure from the perspective of optimization, only one problem is studied. The problem formulation is non-convex and hard to tackle. To conquer this difficulty, a judicious reformulation for such an optimization framework is produced, which leads to the development of an efficient block GS [28] algorithm. Particularly, first, we transform the original optimization problem into another equivalent problem by introducing several auxiliary variables. Next, based on the concept of AO, we propose a custom-made block GS algorithm to optimize the coupled variables, in which both of the gradient projection (GP) method [29] and barrier method [30] are incorporated and available for the construction of an algorithm with high computational efficiency. Finally, we prove the convergence of our proposed algorithm by guaranteeing the KKT [30] optimality under the MIMO multiple ERs scenario.

The rest of this paper is organized as follows. Section II describes the system model and problem statement. Section III and Section IV provide the main contributions of this paper, wherein the efficient resource allocation algorithms are developed for the EHM and WS-EHM problems, respectively. In Section V, we illustrate numerical results to assess the performance of our proposed algorithms. A conclusion is made in Section VI.

*Notations:* The lower-case, boldface lower-case and bold-face upper-case letters are used to denote scalars, vectors and matrices, respectively;  $\text{Tr}(\mathbf{X})$ ,  $|\mathbf{X}|$  ( $\det(\mathbf{X})$ ),  $\mathbf{X}^H$ ,  $(\mathbf{X})^{-1}$ ,  $\|\mathbf{X}\|_F$  and  $\text{rank}(\mathbf{X})$  denote trace, determinant, Hermitian transpose, inverse, Frobenius norm and rank of a matrix  $\mathbf{X}$ ;  $\mathbf{I}$  represents the identity matrix;  $\mathbb{C}^{n \times m}$  denotes the space of  $n \times m$  complex matrices;  $\mathbf{x} \sim \mathcal{CN}(\mu, \Sigma)$  means that the random vector  $\mathbf{x}$  follows a circularly symmetric complex Gaussian (CSCG) distribution with mean  $\mu$  and covariance  $\Sigma$ ;  $\mathbf{X} \succeq \mathbf{0}$  ( $\mathbf{X} \succ \mathbf{0}$ ) means that  $\mathbf{X}$  is positive semidefinite (definite);  $\mathbf{S}_{++}^n$  denotes the symmetric positive definite matrices while  $\mathbb{H}_+^n$  represents  $n \times n$  Hermitian positive semidefinite matrices;  $\mathbb{R}^n$  means  $n$ -dimensional real vectors and  $\mathbb{R}_{++}$  denotes the set of all positive real numbers.

## II. SYSTEM MODEL AND PROBLEM STATEMENT

### A. System Model

Consider a standard three-node MIMO downlink system as shown in Fig. 1(a), which consists of a transmitter, one

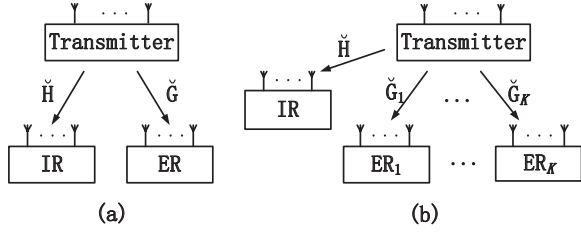


Fig. 1. System model. (a) One multi-antenna ER. (b) Multiple multi-antenna ERs.

IR and one ER. The transmitter who wants to have a secure communication with IR and transmit energy to ER is equipped with  $N_t$  antennas, while the IR and ER are equipped with  $N_i$  and  $N_e$  antennas, resp.. We assume that the ER also acts as a potential eavesdropper to intercept the confidential information intended for IR since the ER is located in the range of service coverage. The channel coefficients from the transmitter to IR and ER are denoted as  $\check{\mathbf{H}} \in \mathbb{C}^{N_t \times N_i}$  and  $\check{\mathbf{G}} \in \mathbb{C}^{N_t \times N_e}$ , resp.. Therefore, the received signals at IR and ER can be expressed as

$$\begin{aligned} \mathbf{y}_i &= \check{\mathbf{H}}^H \mathbf{x} + \mathbf{n}_i, \\ \mathbf{y}_e &= \check{\mathbf{G}}^H \mathbf{x} + \mathbf{n}_e, \end{aligned} \quad (1)$$

resp., where  $\mathbf{x} \in \mathbb{C}^{N_t \times 1}$  is the transmit symbol vector,  $\mathbf{n}_i \sim \mathcal{CN}(\mathbf{0}, \sigma_i^2 \mathbf{I})$  and  $\mathbf{n}_e \sim \mathcal{CN}(\mathbf{0}, \sigma_e^2 \mathbf{I})$  are i.i.d complex additive white Gaussian noises stem from the receiving antennas at the desired receiver. To facilitate SWIPT, the transmit symbol vector is selected as

$$\mathbf{x} = \mathbf{w}_i + \mathbf{w}_e, \quad (2)$$

where  $\mathbf{w}_i \in \mathbb{C}^{N_t \times 1}$  and  $\mathbf{w}_e \in \mathbb{C}^{N_t \times 1}$  are the coded confidential information vector intended for IR and the Gaussian pseudo-random sequences to promote energy transfer, resp.. And  $\mathbf{w}_i$  is assumed to comply with a complex Gaussian distribution  $\mathbf{w}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{W}_I)$ , where  $\mathbf{W}_I \succeq \mathbf{0}$  is the transmit information covariance;  $\mathbf{w}_e$  is modeled as a complex Gaussian pseudo-random vector which is distributed as  $\mathbf{w}_e \sim \mathcal{CN}(\mathbf{0}, \mathbf{W}_E)$ , where  $\mathbf{W}_E \succeq \mathbf{0}$  is the covariance of energy signals. Both of the information and energy signals covariance matrices are the variables to be optimized in our following problems.

### B. Problem Statement

Here, we focus on the EHM problems through the joint design of transmit information and energy covariances under the constraints of quality of service (QoS) requirement on secure communication and given total transmit power budget. It is worth pointing out that the pseudo-random signals which are used for delivering energy are always known at both the transmitter and IR before transmission, hence, the consequent interference can be cancelled at IR as long as the relevant operation is carried out. Accordingly, we consider two types of IRs with and without the ability to execute the interference cancellation about energy signals, and each type corresponds to an EHM problem.

Due to the broadcast nature of wireless communications, assuming unit slot duration, the total energy captured by ER from the received signals  $\mathbf{y}_e$  can be given by [5]–[8]:

$$E(\mathbf{W}_I, \mathbf{W}_E) = \eta \text{Tr} \left[ \check{\mathbf{G}}^H (\mathbf{W}_I + \mathbf{W}_E) \check{\mathbf{G}} \right], \quad (3)$$

where  $0 < \eta \leq 1$ , a constant, is the energy conversion efficiency.

On the other hand, for the purpose of secure communication, given  $(\mathbf{W}_I, \mathbf{W}_E)$ , the achievable secrecy capacity at IR can be depicted as [31]:

$$C_s = C_I(\mathbf{W}_I, \mathbf{W}_E) - C_E(\mathbf{W}_I, \mathbf{W}_E), \quad (4)$$

where  $C_I(\mathbf{W}_I, \mathbf{W}_E)$  is the channel capacity between the transmitter and IR while  $C_E(\mathbf{W}_I, \mathbf{W}_E)$  is the channel capacity between the transmitter and ER. From (4), we can conclude that a perfect secrecy can be ensured when the IR decodes the confidential messages at  $C_s$  bits per channel usage. In other words, at such a decoding rate, the potential eavesdropper (i.e., ER) can retrieve nearly nothing from the messages intended for IR.

In the following context, two EHM problems will be studied for two types of IRs. First, we focus on the case that the IR possesses the capability to cancel the interference from energy signals. On the other hand, of course, interference cancellation can also be performed at the ER when it serves as the potential eavesdropper [6]. The channel capacities  $C_{I1}(\mathbf{W}_I, \mathbf{W}_E)$  and  $C_{E1}(\mathbf{W}_I, \mathbf{W}_E)$  can be written as:

$$\begin{aligned} C_{I1}(\mathbf{W}_I, \mathbf{W}_E) &= \log \left| \mathbf{I} + \frac{1}{\sigma_i^2} \check{\mathbf{H}}^H \mathbf{W}_I \check{\mathbf{H}} \right|, \\ C_{E1}(\mathbf{W}_I, \mathbf{W}_E) &= \log \left| \mathbf{I} + \frac{1}{\sigma_e^2} \check{\mathbf{G}}^H \mathbf{W}_I \check{\mathbf{G}} \right|, \end{aligned} \quad (5)$$

respectively. For notational simplicity, we set  $\mathbf{H} = \frac{1}{\sigma_i} \check{\mathbf{H}}$  and  $\mathbf{G} = \frac{1}{\sigma_e} \check{\mathbf{G}}$ , which are available in the following expression. Then, the first EHM design problem can be formulated as:

$$(P1): \max_{\mathbf{W}_I \succeq \mathbf{0}, \mathbf{W}_E \succeq \mathbf{0}} \sigma_e^2 \eta \text{Tr} [\mathbf{G}^H (\mathbf{W}_I + \mathbf{W}_E) \mathbf{G}] \quad (6a)$$

$$s.t. \underbrace{\log |\mathbf{I} + \mathbf{H}^H \mathbf{W}_I \mathbf{H}| - \log |\mathbf{I} + \mathbf{G}^H \mathbf{W}_I \mathbf{G}|}_{C_{s1}} \geq C_0, \quad (6b)$$

$$\text{Tr}(\mathbf{W}_I + \mathbf{W}_E) \leq P, \quad (6c)$$

where  $C_0 > 0$  is the target secrecy capacity for IR and  $P > 0$  specifies the total power budget<sup>1</sup>.

As mentioned in the introduction, a simplified version of problem (P1) wherein only one optimization variable  $\mathbf{W}_I$  is involved had ever been studied in [9]. However, the solution proposed in [9] is no longer appropriate for problem (P1) because of the additional variable  $\mathbf{W}_E$  which leads to a more complicate problem in terms of optimization. In the next section, we will propose a valid solution for problem (P1) by dividing it into several subproblems.

<sup>1</sup>The constraint (6b) turns out to be concave and the left side of this inequality is positive when the channel is degraded, i.e.,  $\mathbf{H}\mathbf{H}^H - \mathbf{G}\mathbf{G}^H \succeq \mathbf{0}$  holds [9].

Then, we also interest in the EHM problem with the IR not possessing the ability to perform interference cancellation. In this case, the channel capacities  $C_{I2}(\mathbf{W}_I, \mathbf{W}_E)$  and  $C_{E2}(\mathbf{W}_I, \mathbf{W}_E)$  can be given by

$$C_{I2}(\mathbf{W}_I, \mathbf{W}_E) = \log \left| \mathbf{I} + (\mathbf{I} + \mathbf{H}^H \mathbf{W}_E \mathbf{H})^{-1} \mathbf{H}^H \mathbf{W}_I \mathbf{H} \right|, \quad (7a)$$

$$C_{E2}(\mathbf{W}_I, \mathbf{W}_E) = \log \left| \mathbf{I} + \mathbf{G}^H \mathbf{W}_I \mathbf{G} \right|, \quad (7b)$$

resp.. Therefore, the second EHM problem can be stated as:

$$(P2): \max_{\mathbf{W}_I \succeq \mathbf{0}, \mathbf{W}_E \succeq \mathbf{0}} \sigma_e^2 \eta \text{Tr} [\mathbf{G}^H (\mathbf{W}_I + \mathbf{W}_E) \mathbf{G}] \quad (8a)$$

$$s.t. \quad \underbrace{C_{I2}(\mathbf{W}_I, \mathbf{W}_E) - C_{E2}(\mathbf{W}_I, \mathbf{W}_E)}_{C_{s2}} \geq C_0, \quad (8b)$$

$$\text{Tr}(\mathbf{W}_I + \mathbf{W}_E) \leq P. \quad (8c)$$

It can be seen that problem (P2) is a bit more complex than (P1) due to the modified constraint (8b). The solution of problem (P2) will also be demonstrated in the following section.

### III. RESOURCE ALLOCATION SOLUTIONS FOR EHM PROBLEMS

Owing to the fact that both the constraints (6b) and (8b) are non-concave functions with respect to (w.r.t.)  $\mathbf{W}_I$  and  $\mathbf{W}_E$ , without doubt, the optimization problems (P1) and (P2) are both non-convex. Thus, the off-the-shelf tools or methods (such as SeDuMi [26], Lagrange duality [30], ect.) used to deal with convex optimization problems do not apply for them. To solve and derive the valid resource allocation strategies for the EHM problems (P1) and (P2), in this section, we pursue the solutions to these two problems, resp..

#### A. A Water Filling and SVD Based Approach to Problem (P1)

From (6), it is easy to know that the major difficulty in solving problem (P1) lies in the non-concave constraint (6b) which appears to be a fashion of difference of two concave functions. In fact, such a Shannon capacity expression attracts much attention recently and some solutions used for tackling it have been obtained [13], [14], [16], [22]. In [13], the authors assumed a degraded channel model and transmit signals over the positive direction of this channel model to obtain an optimal signaling strategy. Under the assumption, a concave secrecy capacity expression is achieved, and consequently, the corresponding solution. One can refer to [22] for an extension work, where the approach is only limited to the problem concentrate on secrecy rate maximization (SRM). Authors in [14] rewrite the secrecy capacity by a minimax representation through a channel enhancement argument and a clever bounding technique. Whereas, similar to [22], this method works if and only if the secrecy rate is treated as objective. Another common method [16] transforms the secrecy rate expression to an equivalent form by introducing auxiliary variables.

Here, to coordinate with our follow-up further processing, different from aforementioned approaches, we resort to the first order Taylor approximation to reformulate the secrecy rate

constraint (6b). Do expansion at any feasible point  $\mathbf{W}_{I0} \succeq \mathbf{0}$ ,<sup>2</sup> we have [9]

$$\begin{aligned} C_{s1} &\simeq \log \left| \mathbf{I} + \mathbf{H}^H \mathbf{W}_I \mathbf{H} \right| - \log \left| \mathbf{I} + \mathbf{G}^H \mathbf{W}_{I0} \mathbf{G} \right| \\ &\quad - \text{Tr} \left[ \frac{1}{\ln 2} \mathbf{G} (\mathbf{I} + \mathbf{G}^H \mathbf{W}_{I0} \mathbf{G})^{-1} \mathbf{G}^H (\mathbf{W}_I - \mathbf{W}_{I0}) \right] \\ &\triangleq \tilde{C}_{s1}. \end{aligned} \quad (9)$$

After the approximation process on secrecy rate  $C_{s1}$ , it can be seen that the reformulation  $\tilde{C}_{s1}$  is now a concave function w.r.t. the variable  $\mathbf{W}_I$ . Based on this, the EHM problem (P1) can be re-expressed as:

$$\begin{aligned} (\tilde{P1}): \max_{\mathbf{W}_I \succeq \mathbf{0}, \mathbf{W}_E \succeq \mathbf{0}} & \sigma_e^2 \eta \text{Tr} [\mathbf{G}^H (\mathbf{W}_I + \mathbf{W}_E) \mathbf{G}] \\ s.t. & \quad \tilde{C}_{s1} \geq C_0, \\ & \quad \text{Tr}(\mathbf{W}_I + \mathbf{W}_E) \leq P. \end{aligned} \quad (10)$$

Obviously, the representation  $(\tilde{P1})$  is a convex semidefinite programming (SDP) [30] problem, with fixed  $\mathbf{W}_{I0}$ , which can be conveniently and effectively solved by the existing convex optimization tools (e.g. SeDuMi [26] and CVX [27]) in hand in a globally optimal manner. Or, instead, solve its Lagrange dual problem based on the alternative optimization of the original variables and Lagrangian multipliers. However, it is inefficient in terms of computational efficiency, dealing with many SDPs during the update of the initial value  $\mathbf{W}_{I0}$ , no matter which solution is adopted. Therefore, here, we propose a water filling (WF) and SVD based approach to achieve a fast implementation for problem  $(\tilde{P1})$ .

From the construction of problem  $(\tilde{P1})$ , it is easy to learn the independence between variables  $\mathbf{W}_I$  and  $\mathbf{W}_E$ . Hence, by introducing an auxiliary variable  $0 < \alpha \leq 1$ , the optimization problem (10) can be divided into three subproblems

$$(\tilde{P1} - 1): \max_{\mathbf{W}_I \succeq \mathbf{0}} \sigma_e^2 \eta \text{Tr} (\mathbf{G}^H \mathbf{W}_I \mathbf{G}) \quad (11a)$$

$$s.t. \quad \tilde{C}_{s1} \geq C_0, \quad (11b)$$

$$\text{Tr}(\mathbf{W}_I) \leq \alpha P, \quad (11c)$$

$$\begin{aligned} (\tilde{P1} - 2): \max_{\mathbf{W}_E \succeq \mathbf{0}} & \sigma_e^2 \eta \text{Tr} (\mathbf{G}^H \mathbf{W}_E \mathbf{G}) \\ s.t. & \quad \text{Tr}(\mathbf{W}_E) \leq (1 - \alpha) P, \end{aligned} \quad (12)$$

$$\begin{aligned} (\tilde{P1} - 3): \max_{\alpha} & h(\alpha) \\ s.t. & \quad 0 < \alpha \leq 1, \end{aligned} \quad (13)$$

where  $h(\alpha)$  denotes the optimal value of problem  $(\tilde{P1})$ . In fact, above subproblems can be classified into a two-layer optimization problem, where the inner layer contains two convex problems  $(\tilde{P1} - 1)$  and  $(\tilde{P1} - 2)$  with fixed  $\alpha$  while the outer layer is an one-dimensional search problem  $(\tilde{P1} - 3)$  with variable  $\alpha$ .

<sup>2</sup>It is worth to note that the implication  $\tilde{C}_{s1} \geq C_0 \Rightarrow C_{s1} \geq C_0$  holds since  $\tilde{C}_{s1} \leq C_{s1}$  fulfills at any point  $\mathbf{W}_I \succeq \mathbf{0}$  ( $\log |\mathbf{I} + \mathbf{G}^H \mathbf{W}_I \mathbf{G}|$  is concave about  $\mathbf{W}_I$ ), and the equality between  $\tilde{C}_{s1}$  and  $C_{s1}$  satisfies if and only if  $\mathbf{W}_{I0} = \mathbf{W}_I$ , where  $\mathbf{W}_{I0}$  is the update of  $\mathbf{W}_I$  acquired in the last iteration.

1) *Inner Layer Optimization*: Two subproblems are included in the inner layer optimization. First, we focus on subproblem  $(\widetilde{\text{P1}} - 1)$ . Our general idea to solve this subproblem is reformulating it into a mode that can be conveniently tackled by the WF method. Since problem  $(\widetilde{\text{P1}} - 1)$  is convex w.r.t.  $\mathbf{W}_I$  and satisfies the Slaters condition [30], it can be solved by exploiting its dual problem with zero duality gap. Thus, by introducing two non-negative dual variables,  $\lambda$  and  $\mu$ , w.r.t. the secrecy capacity constraint and the transmit power constraint. We have the following Lagrangian and dual function:

$$\mathcal{L}(\lambda, \mu, \mathbf{W}_I) = \sigma_e^2 \eta \text{Tr}(\mathbf{G}^H \mathbf{W}_I \mathbf{G}) + \lambda (\tilde{C}_{s1} - C_0) - \mu [\text{Tr}(\mathbf{W}_I) - \alpha P] \quad (14a)$$

$$d(\lambda, \mu) = \max_{\mathbf{W}_I \succeq \mathbf{0}} \mathcal{L}(\lambda, \mu, \mathbf{W}_I). \quad (14b)$$

The dual problem of  $(\widetilde{\text{P1}} - 1)$ , denoted as  $(\widetilde{\text{P1}} - 1 - D)$ , is defined as  $\min_{\lambda > 0, \mu > 0} d(\lambda, \mu)$ . To solve the dual problem  $(\widetilde{\text{P1}} - 1 - D)$ , first, we maximize the Lagrangian (14a) about  $\mathbf{W}_I$  with fixed  $\lambda$  and  $\mu$ . Then, minimize (14b) to obtain the optimal dual solutions  $\lambda^*$  and  $\mu^*$ . Finally, the dual problem is resolved and the optimal primary solution  $\mathbf{W}_I^*$  that maximizes (14a) is obtained when  $d(\lambda^*, \mu^*)$  is acquired, i.e., the complementary slackness conditions  $\lambda^* (\tilde{C}_{s1}^* - C_0) = 0$  and  $\mu^* [\text{Tr}(\mathbf{W}_I^*) - \alpha P] = 0$  hold [30], where  $\tilde{C}_{s1}^*$  is the optimal secrecy capacity. Moreover, the analytic form of the optimal solution  $\mathbf{W}_I^*$  is worked out and relegated to the following theorem.

**Theorem 3.1:** The optimal solution  $\mathbf{W}_I^*$  of the subproblem  $(\widetilde{\text{P1}} - 1)$  can be formed as follows:

$$\mathbf{W}_I^* = \mathbf{Q}^{-1/2} \mathbf{V} \mathbf{D}^* \mathbf{V}^H \mathbf{Q}^{-1/2}, \quad (15)$$

where  $\mathbf{Q} = \lambda^* \mathbf{G}(\mathbf{I} + \mathbf{G}^H \mathbf{W}_{I0} \mathbf{G})^{-1} \mathbf{G}^H - \sigma_e^2 \eta \mathbf{G} \mathbf{G}^H + \mu^* \mathbf{I} \succ \mathbf{0}$  (the implication  $\mathbf{Q} \succ \mathbf{0}$  must be hold, one can refer to [9] for details),  $\mathbf{V} \in \mathbb{C}^{N_t \times L}$  is derived from the SVD of  $\mathbf{H}^H \mathbf{Q}^{-1/2} = \mathbf{U} \mathbf{\Lambda}^{1/2} \mathbf{V}^H$ , with  $\mathbf{\Lambda}^{1/2} = \text{diag}(\delta_1, \delta_2, \dots, \delta_L)$ ,  $\delta_1 \geq \delta_2 \geq \dots \geq \delta_L > 0$ , and  $\mathbf{D}^* = \text{diag}(p_1^*, \dots, p_L^*)$  with  $p_i^* = (\lambda^* - \frac{1}{\delta_i^2})$ ,  $i = 1, \dots, L$  obtained from the principle of WF power allocation.

*Proof:* Please refer to Appendix A for details.

As mentioned in Appendix A, the ellipsoid method [33] is utilized to optimize the Lagrangian multipliers  $\lambda$  and  $\mu$ . In regarding to the update of these two dual variables, we have the following conclusions.

**Proposition 3.1:** In the process of each iteration of  $\lambda$  and  $\mu$ , we have  $[\lambda^j, \mu^j]^T = [\lambda^{j-1}, \mu^{j-1}]^T - \frac{1}{3} \mathbf{A}^{j-1} \tilde{\mathbf{s}}$ , where  $\lambda^j$  and  $\mu^j$  are the solutions result from the  $j$ th iteration,  $\mathbf{A}^j = \frac{4}{3} (\mathbf{A}^{j-1} - \frac{2}{3} \mathbf{A}^{j-1} \tilde{\mathbf{s}} \tilde{\mathbf{s}}^T \mathbf{A}^{j-1})$  with given initialization  $\mathbf{A}^0 \in \mathbf{S}_{++}^2$  and  $\tilde{\mathbf{s}} = \frac{1}{\sqrt{\mathbf{s}^T \mathbf{A}^{j-1} \mathbf{s}}} \mathbf{s}$  with the subgradient  $\mathbf{s} = [\tilde{C}_{s1}^* - C_0, \alpha P - \text{Tr}(\mathbf{W}_I^*)]^T$ ; the iteration stopping criterion is  $\text{stop} = (\mathbf{s}^T \mathbf{A}^j \mathbf{s})^{1/2} \leq \varepsilon$  ( $\varepsilon > 0$ ), and the iterations will be no more than  $8 \ln(\frac{r l_s}{\varepsilon})$ , where  $r$  is associated with  $\mathbf{A}^0$  while  $l_s$  is associated with the subgradient  $\mathbf{s}$ .

**Algorithm 1** The iterative WF-SVD algorithm for the inner layer optimization.

- 
- 1: **Initialize:**  $j = 1, m = 1, \lambda \geq 0, \mu \geq 0, \varepsilon_1 > 0, \varepsilon_2 > 0, \mathbf{W}_{I0} \succeq \mathbf{0}, \mathbf{W}_I^0 = \mathbf{W}_{I0}$  and  $\mathbf{A} = r^2 \mathbf{I} \succ \mathbf{0}$ , where  $\mathbf{I}$  is a two-dimensional unit matrix;
  - 2: **Repeat**
  - 3:   Set  $\mathbf{A}^0 = \mathbf{A}, \lambda^0 = \lambda, \mu^0 = \mu$ ;
  - 4:   **Repeat**
  - 5:     Compute  $\mathbf{W}_I^*$  through Theorem 3.1;
  - 6:     Calculate the subgradient  $\mathbf{s}$ ;
  - 7:     Update  $\lambda^j, \mu^j$  and  $\mathbf{A}^j$  according to Proposition 3.1;
  - 8:      $j = j + 1$ ;
  - 9:   **Until**  $\text{stop} \leq \varepsilon_1$ ;
  - 10:  $\mathbf{W}_{I0} = \mathbf{W}_I^*, \mathbf{W}_I^m = \mathbf{W}_I^*$  and  $\Delta \mathbf{W}_I = \mathbf{W}_I^m - \mathbf{W}_I^{m-1}$ ;
  - 11:  $m = m + 1$ ;
  - 12: **Until**  $\|\Delta \mathbf{W}_I\|_2 \leq \varepsilon_2$ ;
  - 13: Compute  $\mathbf{W}_E^*$  through Theorem 3.2;
  - 14: **output:**  $\mathbf{W}_I^*$  and  $\mathbf{W}_E^*$ .
- 

*Proof:* See Appendix B.

Next, after the discussion about subproblem  $(\widetilde{\text{P1}} - 1)$ , let's focus on the other subproblem  $(\widetilde{\text{P1}} - 2)$ . Also, similar to the result of previous subproblem, a closed form solution can be derived based on the specific structure of  $(\widetilde{\text{P1}} - 2)$ . Following a parallel mean, which is employed in structural analysis about convex optimization problem, as in [8], the optimal solution  $\mathbf{W}_E^*$  can be presented as follows.

**Theorem 3.2:** The optimal solution  $\mathbf{W}_E^*$  of subproblem  $(\widetilde{\text{P1}} - 2)$  can be given as:

$$\mathbf{W}_E^* = (1 - \alpha) P \mathbf{u}_1 \mathbf{u}_1^H,$$

where  $\mathbf{u}_1 \in \mathbb{C}^{M \times 1}$  represents the first column of the unitary matrix  $\mathbf{U}_G \in \mathbb{C}^{N_t \times N_t}$ , and  $\mathbf{U}_G$  is the result from the SVD of the channel matrix  $\mathbf{G}$  given by  $\mathbf{G} = \mathbf{U}_G \mathbf{\Lambda}_G^{1/2} \mathbf{V}_G^H$ .

*Proof:* The upshot in proving Theorem 3.2 lies in the SVD to both the optimization variable  $\mathbf{W}_E$  and the channel matrix  $\mathbf{G}$ . The proof is not difficult and the details are omitted for brevity. One can refer to [8] for a similar proof.

It is noted that the optimal covariance matrix of the transmit energy signal  $\mathbf{W}_E^*$  is a rank-one matrix and the maximum harvested energy contributed by the energy signals is given by  $E_{\max} = \sigma_e^2 \eta (1 - \alpha) g_1 P$ , where  $g_1$  is the biggest diagonal element w.r.t. to the diagonal matrix  $\mathbf{\Lambda}_G$ . Such a maximum harvested energy reached if and only if the transmit beamforming is performed on the energy signals, which aligns with the most powerful eigenmode of the matrix  $\mathbf{G} \mathbf{G}^H$ . On the basis of above-produced WF and SVD procedures for the inner layer optimization, an iterative WF-SVD algorithm is derived and the details are summarized in Algorithm 1.

*Feasibility assurance about the initial value  $\mathbf{W}_{I0}$ :* The common way to set  $\mathbf{W}_{I0}$  is to initialize it as zeros or a diagonal matrix with the full transmit power uniformly distributed on each element. However, in consideration of the complicated constraints in (11), the initialization may not be feasible

to constraint (11b) when different channel realizations are performed. This will lead to the poor convergence performance and infeasible simulation result of our proposed algorithm. Therefore, to circumvent this defect, the feasibility assurance procedure is necessary to be injected in the determination of the initial value  $\mathbf{W}_{I0}$ . Since the preset  $\mathbf{W}_{I0}$  satisfies the constraint (11c) with fixed  $\alpha$ , our main energy will concentrate on the feasibility check of constraint (11b). To this end, we proceed as follows. Constructing a SCM problem extracted from subproblem  $(\widetilde{\text{P1}} - 1)$

$$\begin{aligned} & \max_{\mathbf{W}_I \succeq \mathbf{0}} \tilde{C}_{s1} \\ \text{s.t.} \quad & \text{Tr}(\mathbf{W}_I) \leq \alpha P. \end{aligned} \quad (16)$$

Given the initialization  $\mathbf{W}_{I0}$ , problem (16) can be solved effectively through the interior point method [30], and the final optimal objective  $\tilde{C}_{s1}^*$ , which must be greater than  $C_0$ , can be achieved based on sequential convex optimizations. As a consequence, the corresponding optimal variable  $\mathbf{W}_I^*$  can be selected as the feasible initial matrix. Alternatively, we can pick out a feasible initialization from the sequential convex operations once the secrecy capacity lower bound  $C_0$  is reached. On the other hand, if the initial matrix  $\mathbf{W}_{I0}$  fulfills the secrecy capacity and total transmit power constraints, no feasibility check procedures are required.

Since both the subproblems  $(\widetilde{\text{P1}} - 1)$  and  $(\widetilde{\text{P1}} - 2)$  are convex, following with the arguments raised in [34], we have the provable guarantee on the convergence of our proposed algorithm as follows:

**Proposition 3.2:** The iterative operations produce non-descending objective values of subproblem  $(\widetilde{\text{P1}} - 1)$ , and it is guaranteed that the optimal solutions  $(\mathbf{W}_I^*$  and  $\mathbf{W}_E^*)$  generated by Algorithm 1 fulfill the KKT optimality conditions of subproblems  $(\widetilde{\text{P1}} - 1)$  and  $(\widetilde{\text{P1}} - 2)$ , resp..

2) *Outer Layer Optimization:* A single-variable optimization problem  $(\widetilde{\text{P1}} - 3)$  is considered in the outer layer part. Since the interval of the single variable  $\alpha$  is bounded and smaller than one, one-dimensional search techniques can be applied to pursue the solution of  $(\widetilde{\text{P1}} - 3)$ . Here, we resort to the efficient golden section search approach which is summarized in Algorithm 2.

So far, we can cope with problem  $(\widetilde{\text{P1}})$  by performing a one-dimensional search during which a series of SDPs are tackled based on the custom-designed algorithm. As a result, the optimal solution  $(\mathbf{W}_I^*, \mathbf{W}_E^*, \alpha^*)$  will be obtained. Moreover, according to [34],  $\mathbf{W}_I^*$  fulfills the KKT conditions of the original problem (P1) though a Taylor linear approximation is implemented about variable  $\mathbf{W}_I$ .

### B. Sequential Convex Optimization to Problem (P2)

It is easy to see that, like problem (P1), problem (P2) is also non-convex owing to the complex non-convex secrecy capacity constraint (8b). However, in our experience, the above mentioned sequential convex optimization framework based on first-order Taylor series expansion is also available for the

**Algorithm 2** The golden section search algorithm for the outer layer optimization.

- 
- 1: **Initialize:**  $\alpha \in (0, 1]$ ,  $a = (\sqrt{5} - 1)/2$ ,  $b = 0$ ,  $c = 1$ ,  $\zeta > 0$ ,  $\alpha_1 = (1 - a)c$  and  $\alpha_2 = ac$
  - 2: Compute  $h(\alpha_1)$  and  $h(\alpha_2)$  based on the results of inner layer optimization;
  - 3: **Repeat**
  - 4:   if  $h(\alpha_1) > h(\alpha_2)$
  - 5:      $c = \alpha_2$ ,  $\alpha_2 = \alpha_1$ ,  $\alpha_1 = c - a(c - b)$ ,
  - 6:     compute  $h(\alpha_1)$  as in step 2;
  - 7:   else
  - 8:      $b = \alpha_1$ ,  $\alpha_1 = \alpha_2$ ,  $\alpha_2 = b + a(c - b)$ ,
  - 9:     compute  $h(\alpha_2)$  as in step 2;
  - 10: **Until**  $|c - b| \leq \zeta$ .
- 

dispose of this problem. In the light of such an idea, we first rewrite the secrecy capacity expression  $C_{s2}$  as:

$$C_{s2} = \log |\mathbf{I} + \mathbf{H}^H (\mathbf{W}_I + \mathbf{W}_E) \mathbf{H}| - \log |\mathbf{I} + \mathbf{H}^H \mathbf{W}_E \mathbf{H}| - \log |\mathbf{I} + \mathbf{G}^H \mathbf{W}_I \mathbf{G}|. \quad (17)$$

Then, for any given  $\mathbf{W}_{I0} \succeq \mathbf{0}$  and  $\mathbf{W}_{E0} \succeq \mathbf{0}$ , Taylor expand the last two terms on the right side of equation (17), we have

$$\begin{aligned} \tilde{C}_{s2} \triangleq & \log |\mathbf{I} + \mathbf{H}^H (\mathbf{W}_I + \mathbf{W}_E) \mathbf{H}| - \log |\mathbf{I} + \mathbf{H}^H \mathbf{W}_{E0} \mathbf{H}| \\ & - \text{Tr} \left[ \frac{1}{\ln 2} \mathbf{H} (\mathbf{I} + \mathbf{H}^H \mathbf{W}_{E0} \mathbf{H})^{-1} \mathbf{H}^H (\mathbf{W}_E - \mathbf{W}_{E0}) \right] \\ & - \log |\mathbf{I} + \mathbf{G}^H \mathbf{W}_{I0} \mathbf{G}| \\ & - \text{Tr} \left[ \frac{1}{\ln 2} \mathbf{G} (\mathbf{I} + \mathbf{G}^H \mathbf{W}_{I0} \mathbf{G})^{-1} \mathbf{G}^H (\mathbf{W}_I - \mathbf{W}_{I0}) \right]. \end{aligned} \quad (18)$$

Based on this approximation, the EHM problem, given any feasible initialization  $(\mathbf{W}_{I0}, \mathbf{W}_{E0})$ , can be reformulated as:

$$\begin{aligned} (\widetilde{\text{P2}}): \quad & \max_{\mathbf{W}_I \succeq \mathbf{0}, \mathbf{W}_E \succeq \mathbf{0}} \sigma_e^2 \eta \text{Tr} [\mathbf{G}^H (\mathbf{W}_I + \mathbf{W}_E) \mathbf{G}] \\ \text{s.t.} \quad & \tilde{C}_{s2} \geq C_0, \\ & \text{Tr}(\mathbf{W}_I + \mathbf{W}_E) \leq P. \end{aligned} \quad (19)$$

It is obvious that the reformulation is convex w.r.t. its variables  $\mathbf{W}_I$  and  $\mathbf{W}_E$ , so it can be globally optimized by the state of the art interior point method with polynomial complexity. Then, in the light of the demonstration in [36], the KKT point, i.e., the optimal solution  $(\mathbf{W}_I^*, \mathbf{W}_E^*)$ , of the original problem (P2) can be determined by performing sequential convex optimization on problem  $(\widetilde{\text{P2}})$ .

## IV. RESOURCE ALLOCATION FOR THE AN-AIDED WS-EHM PROBLEM

In this section, we extend our attention to a generalization of the previously investigated three-node MIMO wiretap channels, where multiple ERs are present and the transmitter is allowed to generate AN [5], [6] for the purpose of promoting efficient energy transfer and guaranteeing secure communication. On this occasion, the WS-EHM problem subject to the constraints of given secrecy rate for IR and limited transmit power of transmitter is considered. Similar to the scenarios assumed in Section II, in this section, we also presume the presence of two types of IRs with and without ability to eliminate the interference incurred by energy signals.

Nevertheless, other than the manner adopted in Section III (deal with each case on its merits), here, we only study one of the two scenarios and custom-design an efficient algorithm for the corresponding optimization problem while the other one is regarded as a simplified version of this one from the perspective of optimization. In order to tackle the resulting optimization problem, many mathematical techniques (such as barrier method, GP method and block GS method) are employed which will be elaborated in the following subsections.

#### A. Signal Model and Problem Formulation

In the secure MIMO multiple ERs communication system shown in Fig. 1(b), the received signals at the IR and  $k$ th ER can be given by

$$\begin{aligned} \mathbf{y}_i &= \widetilde{\mathbf{H}}^H \mathbf{x} + \mathbf{n}_i, \\ \mathbf{y}_{e,k} &= \widetilde{\mathbf{G}}_k^H \mathbf{x} + \mathbf{n}_{e,k}, k \in K, \end{aligned} \quad (20)$$

resp., where  $\mathbf{n}_{e,k} \sim \mathcal{CN}(\mathbf{0}, \sigma_{e,k}^2 \mathbf{I})$  is the complex additive white Gaussian noise and  $K \triangleq \{1, \dots, K\}$ . The transmit signal  $\mathbf{x} \in \mathbb{C}^{N_t \times 1}$  is modeled as:

$$\mathbf{x} = \mathbf{w}_i + \mathbf{w}_e + \mathbf{v},$$

where the newly introduced vector  $\mathbf{v} \in \mathbb{C}^{N_t \times 1}$  is the AN which is generated by the transmitter to facilitate both valid energy transfer and secure communication. The signal vectors  $\mathbf{w}_i$  and  $\mathbf{w}_e$  are assumed to follow the distribution  $\mathbf{w}_i \sim \mathcal{CN}(\mathbf{0}, \mathbf{W}_I)$  and  $\mathbf{w}_e \sim \mathcal{CN}(\mathbf{0}, \mathbf{W}_E)$ , resp., as before, while  $\mathbf{v}$  is modeled as a complex Gaussian vector with  $\mathbf{v} \sim \mathcal{CN}(\mathbf{0}, \mathbf{V})$ .

Based on the setup above, considering the case that the IR is unable to eliminate the interference from energy signals, set  $\mathbf{H} = \frac{1}{\sigma_i} \widetilde{\mathbf{H}}$  and  $\mathbf{G}_k = \frac{1}{\sigma_{e,k}} \widetilde{\mathbf{G}}_k, \forall k \in K$ , we formulate the AN aided WS-EHM problem as follows:

$$(P3): \max_{\mathbf{W}_I \succeq \mathbf{0}, \mathbf{W}_E \succeq \mathbf{0}, \mathbf{V} \succeq \mathbf{0}} E_{WS}(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V}) \quad (21a)$$

$$s.t. \quad R_s \geq R_0 \ln 2, \quad (21b)$$

$$\text{Tr}(\mathbf{W}_I + \mathbf{W}_E + \mathbf{V}) \leq P, \quad (21c)$$

where  $E_{WS}(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V}) = \sum_{k=1}^K \sigma_{e,k}^2 \mu_k \eta_k \text{Tr}[\mathbf{G}_k^H (\mathbf{W}_I + \mathbf{W}_E + \mathbf{V}) \mathbf{G}_k]$ ,  $\mu_k \geq 0$  is the energy weight for the  $k$ th ER,  $0 < \eta_k \leq 1$  is the energy conversion efficiency and we have [15], [31]

$$\begin{aligned} R_s &= \ln \underbrace{\left| \mathbf{I} + [\mathbf{I} + \mathbf{H}^H (\mathbf{W}_E + \mathbf{V}) \mathbf{H}]^{-1} \mathbf{H}^H \mathbf{W}_I \mathbf{H} \right|}_{C_I(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V})} \\ &\quad - \max_{k \in K} \ln \underbrace{\left| \mathbf{I} + [\mathbf{I} + \mathbf{G}_k^H \mathbf{V} \mathbf{G}_k]^{-1} \mathbf{G}_k^H \mathbf{W}_I \mathbf{G}_k \right|}_{C_{E,k}(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V})}. \end{aligned} \quad (22)$$

The formulated WS-EHM problem (P3) is much more complex than the EHM problems mentioned in previous sections. In the following subsections, we will show how problem (P3) can be solved by our Tailor-made algorithm.

#### B. Barrier Reformulation and GP Based Block GS Approach to Problem (P3)

It can be seen that the secrecy rate constraint (21b) is non-convex and it is difficult to deal with it straightly. Therefore, we need the aid of the following lemma.

**Lemma 4.1:** Let  $N$  be any integer and matrix  $\mathbf{E} \in \mathbb{C}^{N \times N}$  is arbitrary such that  $\mathbf{E} \succ \mathbf{0}$ . Consider the function  $f(\mathbf{T}) = \ln |\mathbf{T}| - \text{Tr}(\mathbf{TE}) + N$ . Then,

$$\ln |\mathbf{E}^{-1}| = \max_{\mathbf{T} \in \mathbb{C}^{N \times N}, \mathbf{T} \succ \mathbf{0}} f(\mathbf{T}),$$

with the optimal solution  $\mathbf{T}^* = \mathbf{E}^{-1}$ .

Using Lemma 4.1, by setting  $\mathbf{E} = \mathbf{I} + \mathbf{H}^H (\mathbf{W}_E + \mathbf{V}) \mathbf{H}$ ,  $C_I(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V})$  can be rewritten as

$$\begin{aligned} C_I &= \ln \left| [\mathbf{I} + \mathbf{H}^H (\mathbf{W}_E + \mathbf{V}) \mathbf{H}]^{-1} \right| \\ &\quad + \ln \left| \mathbf{I} + \mathbf{H}^H (\mathbf{W}_I + \mathbf{W}_E + \mathbf{V}) \mathbf{H} \right| \\ &= \max_{\mathbf{T}_0 \succeq \mathbf{0}} \ln |\mathbf{T}_0| - \text{Tr} \{ \mathbf{T}_0 [\mathbf{I} + \mathbf{H}^H (\mathbf{W}_E + \mathbf{V}) \mathbf{H}] \} \\ &\quad + N_i + \ln \left| \mathbf{I} + \mathbf{H}^H (\mathbf{W}_I + \mathbf{W}_E + \mathbf{V}) \mathbf{H} \right| \\ &= \max_{\mathbf{T}_0 \succeq \mathbf{0}} \theta_I(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V}, \mathbf{T}_0), \end{aligned} \quad (23)$$

where  $\mathbf{T}_0 \in \mathbb{H}_+^{N_i}$  and  $C_I(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V})$  is simplified to  $C_I$  for the convenience of description. Utilizing Lemma 4.1 again, set  $\mathbf{E}_k = \mathbf{I} + \mathbf{G}_k^H (\mathbf{W}_I + \mathbf{W}_E + \mathbf{V}) \mathbf{G}_k, \forall k \in K$ , we can rewrite  $C_{E,k}(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V})$  as

$$\begin{aligned} C_{E,k} &= \ln |\mathbf{I} + \mathbf{G}_k^H \mathbf{V} \mathbf{G}_k|^{-1} + \ln |\mathbf{I} + \mathbf{G}_k^H (\mathbf{W}_I + \mathbf{V}) \mathbf{G}_k| \\ &= \min_{\mathbf{T}_k \succeq \mathbf{0}} -\ln |\mathbf{I} + \mathbf{G}_k^H \mathbf{V} \mathbf{G}_k| - \ln |\mathbf{T}_k| \\ &\quad + \text{Tr} \{ \mathbf{T}_k [\mathbf{I} + \mathbf{G}_k^H (\mathbf{W}_I + \mathbf{V}) \mathbf{G}_k] \} - N_{e,k} \\ &= \min_{\mathbf{T}_k \succeq \mathbf{0}} \theta_{E,k}(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V}, \mathbf{T}_k), \forall k \in K, \end{aligned} \quad (24)$$

where  $\mathbf{T}_k \in \mathbb{H}_+^{N_e}$  and  $C_{E,k} = C_{E,k}(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V})$ . By substituting (23) and (24) into (22), we reformulate problem (P3) as

$$\begin{aligned} (\widetilde{P3}): \quad & \max_{\mathbf{W}_I \succeq \mathbf{0}, \mathbf{W}_E \succeq \mathbf{0}, \mathbf{V} \succeq \mathbf{0}} E_{WS}(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V}) \\ s.t. \quad & \max_{\mathbf{T}_0 \succeq \mathbf{0}, \mathbf{T}_k \succeq \mathbf{0}} \theta_I(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V}, \mathbf{T}_0) \\ & - \theta_{E,k}(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V}, \mathbf{T}_k) \geq R_0 \ln 2, \forall k \in K, \\ & \text{Tr}(\mathbf{W}_I + \mathbf{W}_E + \mathbf{V}) \leq P. \end{aligned} \quad (25)$$

From (25), we can see that a sequence of auxiliary variables  $\{\mathbf{T}_k\}_{k=0}^K$  are introduced in reformulation  $(\widetilde{P3})$ . Considering the existence of coupled variables and the feasible set is closed, nonempty and convex, we employ a block GS method [30] consists in solving two or three subproblems iteratively to settle problem  $(\widetilde{P3})$ . Specifically, let  $(\mathbf{W}_I^n, \mathbf{W}_E^n, \mathbf{V}^n, \{\mathbf{T}_k^n\}_{k=0}^K)$  be the iterate at the  $n$ th iteration, each iteration of the block GS method consists of following subproblems:

$$(\widetilde{P3}-1): \mathbf{T}_0^n = \arg \max_{\mathbf{T}_0 \succeq \mathbf{0}} \theta_I(\mathbf{W}_I^{n-1}, \mathbf{W}_E^{n-1}, \mathbf{V}^{n-1}, \mathbf{T}_0), \quad (26)$$

$$\begin{aligned} (\widetilde{P3}-2): \quad & \mathbf{T}_k^n = \\ & \arg \max_{\mathbf{T}_k \succeq \mathbf{0}} \theta_{E,k}(\mathbf{W}_I^{n-1}, \mathbf{W}_E^{n-1}, \mathbf{V}^{n-1}, \mathbf{T}_k), \forall k \in K, \end{aligned} \quad (27)$$

$$\begin{aligned} (\widetilde{\text{P3}} - 3) : (\mathbf{W}_I^n, \mathbf{W}_E^n, \mathbf{V}^n) = \\ \arg \max_{(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V}) \in \Theta} E_{\text{WS}}(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V}), \end{aligned} \quad (28)$$

where  $\Theta = \{(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V}) \mid \theta_I(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V}, \mathbf{T}_0^n) - \theta_{E,k}(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V}, \mathbf{T}_k^n) \geq R_0 \ln 2, \forall k \in K, \text{Tr}(\mathbf{W}_I + \mathbf{W}_E + \mathbf{V}) \leq P, \mathbf{W}_I \succeq \mathbf{0}, \mathbf{W}_E \succeq \mathbf{0}, \mathbf{V} \succeq \mathbf{0}\}$ , for  $n = 1, 2, \dots$ . In the light of Lemma 4.1, given the value of  $(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V})$  in the previous iteration, the solutions of subproblems  $(\widetilde{\text{P3}} - 1)$  and  $(\widetilde{\text{P3}} - 2)$  in the next iteration can be easily given by

$$\mathbf{T}_0^n = [\mathbf{I} + \mathbf{H}^H (\mathbf{W}_E^{n-1} + \mathbf{V}^{n-1}) \mathbf{H}]^{-1}, \quad (29)$$

$$\mathbf{T}_k^n = [\mathbf{I} + \mathbf{G}_k^H (\mathbf{W}_I^{n-1} + \mathbf{V}^{n-1}) \mathbf{G}_k]^{-1}, \forall k \in K. \quad (30)$$

For subproblem  $(\widetilde{\text{P3}} - 3)$ , since it is convex w.r.t. variables  $\mathbf{W}_I, \mathbf{W}_E$  and  $\mathbf{V}$  when  $\mathbf{T}_0^n$  and  $\{\mathbf{T}_k^n\}_{k=1}^K$  are fixed, the state of the art solver, such as the SeDuMi, can be used to solve this problem. Therefore, with such a conventional mind, we will solve a sequence of SDPs during the iteration. Here, instead, we put forward an alternative method in solving  $(\widetilde{\text{P3}} - 3)$  with much higher computational efficiency than the general method (cf. Table II). We will first perform a logarithmic barrier reformulation of subproblem  $(\widetilde{\text{P3}} - 3)$ , and then design a high-efficiency algorithm for the reformulated problem on the basis of GP method.

Two inequality constraints are included in  $(\widetilde{\text{P3}} - 3)$ , to facilitate the subsequent process with GP method, we need to get rid of the secrecy rate constraint while no substantial change will be occurred to our problem. To this end, we resort to such an indicator function

$$I_+(x_k^n) = \begin{cases} 0, & x_k^n \geq 0, \\ -\infty, & x_k^n < 0, \end{cases} \quad (31)$$

where  $I_+ : \mathbb{R} \rightarrow \mathbb{R}$  is the indicator function for the non-negative reals and  $x_k^n = x_k(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V}, \{\mathbf{T}_k^n\}_{k=0}^K) = \theta_I(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V}, \mathbf{T}_0^n) - \theta_{E,k}(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V}, \mathbf{T}_k^n) - R_0 \ln 2, \forall k \in K$ . With this function, the secrecy rate constraint in subproblem  $(\widetilde{\text{P3}} - 3)$  can be removed to the objective and acts implicitly. When  $x_k^n \geq 0$  holds during the iteration, the solution of the reformulation will be consistent with our original problem, if not, the objective value will be punished to negative infinity. However, it is obvious that the piecewise function in (31) is non-differentiable. As a consequence, we introduce the logarithmic barrier function to approximate the indicator function, which can be depicted as:

$$\widehat{I}_+(x_k^n) = \frac{1}{t} \ln(x_k^n), \text{dom} \widehat{I}_+ = \mathbb{R}_{++}, \quad (32)$$

where  $t > 0$  is a parameter that sets the accuracy of the approximation. Note that  $x_k^n$  is concave w.r.t.  $(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V})$

<sup>3</sup>The accuracy of the approximation is determined by the parameter  $t$ . As  $t$  increase,  $\widehat{I}_+$  will be closer to  $I_+$ , i.e.,  $\widehat{I}_+(x_k^n)$  will be closer to zero when  $x_k^n \geq 0$  is satisfied. The difference between both functions disappear once the value of  $t$  tend to be infinite, i.e.,  $t \rightarrow \infty$ .

and the logarithmic operation in (32) is concave and non-decreasing, thus, according to the concept of convexity preserving operation [30], we conclude that the logarithmic barrier function  $\widehat{I}_+(x_k^n)$  is concave w.r.t.  $(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V})$ .

Making (32) in the objective of subproblem  $(\widetilde{\text{P3}} - 3)$ , we get the reformulation

$$\begin{aligned} (\widetilde{\text{P3}} - 3A) : (\mathbf{W}_I^n, \mathbf{W}_E^n, \mathbf{V}^n) = \\ \arg \max_{(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V}) \in \Xi} E_{\text{WS}}(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V}) + \frac{1}{t} \sum_{k=1}^K \ln(x_k^n) \end{aligned} \quad (33)$$

where  $\Xi = \{(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V}) \mid \text{Tr}(\mathbf{W}_I + \mathbf{W}_E + \mathbf{V}) \leq P, \mathbf{W}_I \succeq \mathbf{0}, \mathbf{W}_E \succeq \mathbf{0}, \mathbf{V} \succeq \mathbf{0}\}$  and the Cumulative-Sum operation on the logarithmic barrier function is due to the existence of  $K$  secrecy rate constraints.

To provide a solver to problem  $(\widetilde{\text{P3}} - 3A)$ , we turn to the GP method which possesses low computational cost by exploiting the problem structure. Let  $g(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V}) = E_{\text{WS}}(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V}) + \frac{1}{t} \sum_{k=1}^K \ln(x_k^n)$ . The GP method applied to problem  $(\widetilde{\text{P3}} - 3A)$  is defined as follows [29]:

$$\begin{aligned} (\widehat{\mathbf{W}}_I^{n,m+1}, \widehat{\mathbf{W}}_E^{n,m+1}, \widehat{\mathbf{V}}^{n,m+1}) \\ = \text{Proj}_{\Xi}[(\mathbf{W}_I^{n,m}, \mathbf{W}_E^{n,m}, \mathbf{V}^{n,m}) \\ + q_1 (\nabla_{\mathbf{W}_I} g^{n,m}, \nabla_{\mathbf{W}_E} g^{n,m}, \nabla_{\mathbf{V}} g^{n,m}), \end{aligned} \quad (34)$$

$$\begin{aligned} (\mathbf{W}_I^{n,m+1}, \mathbf{W}_E^{n,m+1}, \mathbf{V}^{n,m+1}) = (\mathbf{W}_I^{n,m}, \mathbf{W}_E^{n,m}, \mathbf{V}^{n,m}) + \\ q_2 ((\widehat{\mathbf{W}}_I^{n,m+1}, \widehat{\mathbf{W}}_E^{n,m+1}, \widehat{\mathbf{V}}^{n,m+1}) - (\mathbf{W}_I^{n,m}, \mathbf{W}_E^{n,m}, \mathbf{V}^{n,m})) \end{aligned} \quad (35)$$

where  $(\widehat{\mathbf{W}}_I^{n,m+1}, \widehat{\mathbf{W}}_E^{n,m+1}, \widehat{\mathbf{V}}^{n,m+1})$  is the introduced auxiliary variable set during the iteration  $m+1$  of GP method;  $\text{Proj}_{\Xi}$  denotes the projection on feasible set  $\Xi$  and its analytical expression is depicted as:

$$\begin{aligned} \text{Proj}_{\Xi}(\bar{\mathbf{W}}_I, \bar{\mathbf{W}}_E, \bar{\mathbf{V}}) = \arg \min_{(\widehat{\mathbf{W}}_I, \widehat{\mathbf{W}}_E, \widehat{\mathbf{V}}) \in \Xi} \left\| \widehat{\mathbf{W}}_I - \bar{\mathbf{W}}_I \right\|_F^2 \\ + \left\| \widehat{\mathbf{W}}_E - \bar{\mathbf{W}}_E \right\|_F^2 + \left\| \widehat{\mathbf{V}} - \bar{\mathbf{V}} \right\|_F^2; \end{aligned} \quad (36)$$

$q_1 > 0$  and  $0 < q_2 \leq 1$  are the step sizes, and one can refer to the Armijo rule [29] for the details about the step size selection;  $g$ , a shorter version, denotes  $g(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V})$  and  $(\nabla_{\mathbf{W}_I} g^{n,m}, \nabla_{\mathbf{W}_E} g^{n,m}, \nabla_{\mathbf{V}} g^{n,m})$  denotes the gradient of  $g(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V})$  w.r.t.  $(\mathbf{W}_I^{n,m}, \mathbf{W}_E^{n,m}, \mathbf{V}^{n,m})$ , which is given particularly in (37) on the top of next page, where  $\Omega^{n,m} = \mathbf{H}[\mathbf{I} + \mathbf{H}^H (\mathbf{W}_I^{n,m} + \mathbf{W}_E^{n,m} + \mathbf{V}^{n,m}) \mathbf{H}]^{-1} \mathbf{H}^H$  and  $\Psi^{n,m} = \mathbf{G}_k[\mathbf{I} + \mathbf{G}_k^H \mathbf{V}^{n,m} \mathbf{G}_k]^{-1} \mathbf{G}_k^H, \forall k \in K$ .

The core of the GP method lies in the projection process in (36), by taking full advantage of the structure of feasible region  $\Xi$ , we can implement the projection effectively.

**Proposition 4.1:** Similar to the skills, SVD and EVD, applied in dealing with subproblem  $(\widetilde{\text{P1}} - 2)$ , carry out EVD of each known quantity in (36), i.e.,  $\bar{\mathbf{W}}_I = \mathbf{U}_{\mathbf{W}_I} \bar{\Lambda}_{\mathbf{W}_I} \mathbf{U}_{\mathbf{W}_I}^H$ ,  $\bar{\mathbf{W}}_E = \mathbf{U}_{\mathbf{W}_E} \bar{\Lambda}_{\mathbf{W}_E} \mathbf{U}_{\mathbf{W}_E}^H$  and  $\bar{\mathbf{V}} = \mathbf{U}_{\mathbf{V}} \bar{\Lambda}_{\mathbf{V}} \mathbf{U}_{\mathbf{V}}^H$ . Then, we



$$\begin{cases} \nabla_{\mathbf{W}_I} g^{n,m} = \sum_{k=1}^K \sigma_{e,k}^2 \mu_k \eta_k \mathbf{G}_k \mathbf{G}_k^H + \frac{1}{t} \sum_{k=1}^K \frac{1}{x_k^n} (\boldsymbol{\Omega}^{n,m} - \mathbf{G}_k \mathbf{T}_k^n \mathbf{G}_k^H), \\ \nabla_{\mathbf{W}_E} g^{n,m} = \sum_{k=1}^K \sigma_{e,k}^2 \mu_k \eta_k \mathbf{G}_k \mathbf{G}_k^H + \frac{1}{t} \sum_{k=1}^K \frac{1}{x_k^n} (-\mathbf{H} \mathbf{T}_0^n \mathbf{H}^H + \boldsymbol{\Omega}^{n,m}), \\ \nabla_{\mathbf{V}} g^{n,m} = \sum_{k=1}^K \sigma_{e,k}^2 \mu_k \eta_k \mathbf{G}_k \mathbf{G}_k^H + \frac{1}{t} \sum_{k=1}^K \frac{1}{x_k^n} (-\mathbf{H} \mathbf{T}_0^n \mathbf{H}^H + \boldsymbol{\Omega}^{n,m} + \boldsymbol{\Psi}^{n,m} - \mathbf{G}_k \mathbf{T}_k^n \mathbf{G}_k^H), \end{cases} \quad (37)$$

have

$$\begin{aligned} \text{Proj}_{\Xi}(\bar{\mathbf{W}}_I, \bar{\mathbf{W}}_E, \bar{\mathbf{V}}) &= (\widehat{\mathbf{W}}_I, \widehat{\mathbf{W}}_E, \widehat{\mathbf{V}}) \\ &= (\mathbf{U}_{\mathbf{W}_I} \widehat{\boldsymbol{\Lambda}}_{\mathbf{W}_I} \mathbf{U}_{\mathbf{W}_I}^H, \mathbf{U}_{\mathbf{W}_E} \widehat{\boldsymbol{\Lambda}}_{\mathbf{W}_E} \mathbf{U}_{\mathbf{W}_E}^H, \mathbf{U}_{\mathbf{V}} \widehat{\boldsymbol{\Lambda}}_{\mathbf{V}} \mathbf{U}_{\mathbf{V}}^H). \end{aligned}$$

Motivated by this, we now turn our attention to the calculation of  $(\widehat{\boldsymbol{\Lambda}}_{\mathbf{W}_I}, \widehat{\boldsymbol{\Lambda}}_{\mathbf{W}_E}, \widehat{\boldsymbol{\Lambda}}_{\mathbf{V}})$  instead of  $(\widehat{\mathbf{W}}_I, \widehat{\mathbf{W}}_E, \widehat{\mathbf{V}})$ .

$$(\widehat{\boldsymbol{\Lambda}}_{\mathbf{W}_I}^*, \widehat{\boldsymbol{\Lambda}}_{\mathbf{W}_E}^*, \widehat{\boldsymbol{\Lambda}}_{\mathbf{V}}^*) = \arg \min_{(\widehat{\boldsymbol{\Lambda}}_{\mathbf{W}_I}, \widehat{\boldsymbol{\Lambda}}_{\mathbf{W}_E}, \widehat{\boldsymbol{\Lambda}}_{\mathbf{V}}) \in \Xi} \|\widehat{\boldsymbol{\Lambda}} - \bar{\boldsymbol{\Lambda}}\|_F^2 \quad (38)$$

where  $\widehat{\boldsymbol{\Lambda}} = (\text{diag}(\widehat{\boldsymbol{\Lambda}}_{\mathbf{W}_I}), \text{diag}(\widehat{\boldsymbol{\Lambda}}_{\mathbf{W}_E}), \text{diag}(\widehat{\boldsymbol{\Lambda}}_{\mathbf{V}}))$  and  $\bar{\boldsymbol{\Lambda}} = (\text{diag}(\bar{\boldsymbol{\Lambda}}_{\mathbf{W}_I}), \text{diag}(\bar{\boldsymbol{\Lambda}}_{\mathbf{W}_E}), \text{diag}(\bar{\boldsymbol{\Lambda}}_{\mathbf{V}}))$ .

*Proof:* First of all, by writing  $\bar{\mathbf{W}}_I$ ,  $\bar{\mathbf{W}}_E$  and  $\bar{\mathbf{V}}$  in their EVD forms, we obtain the equivalent transformation of (36) which is shown in (39), where  $\widehat{\boldsymbol{\Lambda}}_{\mathbf{W}_I}$ ,  $\widehat{\boldsymbol{\Lambda}}_{\mathbf{W}_E}$  and  $\widehat{\boldsymbol{\Lambda}}_{\mathbf{V}}$  are defined as  $\mathbf{U}_{\mathbf{W}_I}^H \bar{\mathbf{W}}_I \mathbf{U}_{\mathbf{W}_I}$ ,  $\mathbf{U}_{\mathbf{W}_E}^H \bar{\mathbf{W}}_E \mathbf{U}_{\mathbf{W}_E}$  and  $\mathbf{U}_{\mathbf{V}}^H \bar{\mathbf{V}} \mathbf{U}_{\mathbf{V}}$ , resp.. Our next endeavor is to certify that the optimal solution  $(\widehat{\boldsymbol{\Lambda}}_{\mathbf{W}_I}^*, \widehat{\boldsymbol{\Lambda}}_{\mathbf{W}_E}^*, \widehat{\boldsymbol{\Lambda}}_{\mathbf{V}}^*)$  of (38) must be diagonal. Assuming an optimal solution  $(\widehat{\boldsymbol{\Lambda}}'_{\mathbf{W}_I}, \widehat{\boldsymbol{\Lambda}}'_{\mathbf{W}_E}, \widehat{\boldsymbol{\Lambda}}'_{\mathbf{V}})$  which is not diagonal and the resulting objective value is  $y'$ . According to (39), it is easy to learn that there must have a diagonal counterpart which achieves a smaller objective than  $y'$  in the feasible zone  $\Xi$ . Noticing such a contradiction, we conclude that (36) is equivalent to its simplified version (38).

In accordance with the argument in [35], the optimal solution of (38) can be given by

$$\begin{aligned} \widehat{\boldsymbol{\Lambda}}_{\mathbf{W}_I}^* &= (\bar{\boldsymbol{\Lambda}}_{\mathbf{W}_I} - \rho^* \mathbf{I})^+, \widehat{\boldsymbol{\Lambda}}_{\mathbf{W}_E}^* = (\bar{\boldsymbol{\Lambda}}_{\mathbf{W}_E} - \rho^* \mathbf{I})^+, \\ \widehat{\boldsymbol{\Lambda}}_{\mathbf{V}}^* &= (\bar{\boldsymbol{\Lambda}}_{\mathbf{V}} - \rho^* \mathbf{I})^+, \end{aligned} \quad (40)$$

where  $\rho^*$  can be calculated efficiently through the Algorithm 1 in [35] and  $(x)^+ \triangleq \max(0, x)$ .

So far, we accomplish the derivation of our custom-made barrier-GP method for subproblem  $(\widetilde{\text{P3}} - 3)$ . The details of barrier-GP method together with the whole block GS process are summarized in Algorithm 3.

Like the measures that we ever took in Algorithm 1, here, the feasibility assurance procedure of the initial values  $\mathbf{W}_I^0$ ,  $\mathbf{W}_E^0$  and  $\mathbf{V}^0$  is also necessary during each channel realization. It should be noted that in Algorithm 3, the parameter  $t$  is updated (increased) by a factor  $\mu_t > 1$ , under the guidance of central path which converges to the optimal point as  $t \rightarrow \infty$  [30]. Since the value of  $t$  can be updated at each iteration and tends to be infinite ( $\xi_2$  is infinitesimal), the reformulation  $(\widetilde{\text{P3}} - 3\text{A})$  can be deemed to be equivalent to  $(\widetilde{\text{P3}} - 3)$ . Therefore, the resulting solution will be optimal to problem  $(\widetilde{\text{P3}})$  as well as problem (P3).

**Algorithm 3** The barrier-GP based block GS algorithm for WS-EHM (P3).

- 1: **Initialize:**  $n = 0$ ,  $q_1 > 0$ ,  $\mu_t > 1$ ,  $\xi_1 > 0$ ,  $\xi_2 > 0$ ,  $\xi_3 > 0$ ,  $\mathbf{W}_I^0 \succeq \mathbf{0}$ ,  $\mathbf{W}_E^0 \succeq \mathbf{0}$ ,  $\mathbf{V}^0 \succeq \mathbf{0}$ ;
- 2: **Repeat**
- 3:    $n = n + 1$ ,  $t = 1$ ;
- 4:   Calculate  $\mathbf{T}_0^n$  and  $\{\mathbf{T}_k^n\}_{k=1}^K$  according to (26) and (27), resp.;
- 5:   Set  $(\mathbf{W}_I^{n,0}, \mathbf{W}_E^{n,0}, \mathbf{V}^{n,0}) = (\mathbf{W}_I^{n-1}, \mathbf{W}_E^{n-1}, \mathbf{V}^{n-1})$ ;
- 6:   **Repeat**
- 7:      $m = 0$ ;
- 8:     **Repeat**
- 9:       Compute  $(\nabla_{\mathbf{W}_I} g^{n,m}, \nabla_{\mathbf{W}_E} g^{n,m}, \nabla_{\mathbf{V}} g^{n,m})$  in line with (37);
- 10:       Compute  $(\widehat{\mathbf{W}}_I^{n,m+1}, \widehat{\mathbf{W}}_E^{n,m+1}, \widehat{\mathbf{V}}^{n,m+1})$  according to Proposition 4.1 and (40);
- 11:       Compute  $(\mathbf{W}_I^{n,m}, \mathbf{W}_E^{n,m}, \mathbf{V}^{n,m})$  according to (35), where the value of  $q_2$  is determined by the Armijo rule in [29];
- 12:        $m = m + 1$ ;
- 13:       **Until**  $|(g^{n,m} - g^{n,m-1})/g^{n,m-1}| < \xi_1$
- 14:       Update  $t = \mu_t t$ ;
- 15:       **Until**  $K/t < \xi_2$
- 16:       Update  $\mathbf{W}_I^n = \mathbf{W}_I^{n,m}$ ,  $\mathbf{W}_E^n = \mathbf{W}_E^{n,m}$ ,  $\mathbf{V}^n = \mathbf{V}^{n,m}$ ;
- 17:       **Until**  $|\mathbf{E}_{\text{WS}}^n - \mathbf{E}_{\text{WS}}^{n-1}| < \xi_3$
- 18:    $\mathbf{W}_I^* = \mathbf{W}_I^n$ ,  $\mathbf{W}_E^* = \mathbf{W}_E^n$ ,  $\mathbf{V}^* = \mathbf{V}^n$ ;
- 19: **Output:**  $\mathbf{W}_I^*$ ,  $\mathbf{W}_E^*$  and  $\mathbf{V}^*$ .

**Proposition 4.2:** The monotonically non-decreasing objective values are produced from our proposed barrier-GP based block GS algorithm. Moreover, every limit point  $(\mathbf{W}_I^*, \mathbf{W}_E^*, \mathbf{V}^*)$  of the iterations generated by the block GS process is a KKT point of problem (P3).

*Proof:* Please refer to Appendix C.

At this point, we finished the study of AN-aided WS-EHM problem for the case that the IR is not able to eliminate the interference from the energy signals. Since the whole process presented above is still available for the problem dropping  $\mathbf{W}_E$  from (22), the research of the case in which the IR is able to remove the interference from the energy signals is omitted here.

## V. NUMERICAL RESULTS

In this section, we provide numerical results to validate the performance of our proposed algorithms for EHM and WS-EHM. In each simulation trial, the distance-dependent path

$$\begin{aligned}
& \min_{(\widehat{\mathbf{W}}_I, \widehat{\mathbf{W}}_E, \widehat{\mathbf{V}}) \in \Xi} \left\| \widehat{\mathbf{W}}_I - \mathbf{U}_{\mathbf{W}_I} \bar{\mathbf{A}}_{\mathbf{W}_I} \mathbf{U}_{\mathbf{W}_I}^H \right\|_F^2 + \left\| \widehat{\mathbf{W}}_E - \mathbf{U}_{\mathbf{W}_E} \bar{\mathbf{A}}_{\mathbf{W}_E} \mathbf{U}_{\mathbf{W}_E}^H \right\|_F^2 + \left\| \widehat{\mathbf{V}} - \mathbf{U}_{\mathbf{V}} \bar{\mathbf{A}}_{\mathbf{V}} \mathbf{U}_{\mathbf{V}}^H \right\|_F^2 \\
& \Leftrightarrow \min_{(\widehat{\mathbf{W}}_I, \widehat{\mathbf{W}}_E, \widehat{\mathbf{V}}) \in \Xi} \left\| \mathbf{U}_{\mathbf{W}_I}^H \widehat{\mathbf{W}}_I \mathbf{U}_{\mathbf{W}_I} - \bar{\mathbf{A}}_{\mathbf{W}_I} \right\|_F^2 + \left\| \mathbf{U}_{\mathbf{W}_E}^H \widehat{\mathbf{W}}_E \mathbf{U}_{\mathbf{W}_E} - \bar{\mathbf{A}}_{\mathbf{W}_E} \right\|_F^2 + \left\| \mathbf{U}_{\mathbf{V}}^H \widehat{\mathbf{V}} \mathbf{U}_{\mathbf{V}} - \bar{\mathbf{A}}_{\mathbf{V}} \right\|_F^2 \\
& \Leftrightarrow \min_{(\widehat{\mathbf{A}}_{\mathbf{W}_I}, \widehat{\mathbf{A}}_{\mathbf{W}_E}, \widehat{\mathbf{A}}_{\mathbf{V}}) \in \Xi} \left\| \widehat{\mathbf{A}}_{\mathbf{W}_I} - \bar{\mathbf{A}}_{\mathbf{W}_I} \right\|_F^2 + \left\| \widehat{\mathbf{A}}_{\mathbf{W}_E} - \bar{\mathbf{A}}_{\mathbf{W}_E} \right\|_F^2 + \left\| \widehat{\mathbf{A}}_{\mathbf{V}} - \bar{\mathbf{A}}_{\mathbf{V}} \right\|_F^2,
\end{aligned} \tag{39}$$

loss model is given by

$$\tilde{L} = A_0 \left( \frac{d}{d_0} \right)^{-\gamma}, \quad d > d_0,$$

where  $A_0=1$ ,  $d$  denotes the distance between the transmitter and the associated receiver (IR or ER),  $d_0$  is a reference distance set to be 1m, and  $\gamma$  is the path loss exponent set at 3.

#### A. One ER EHM

In this subsection, we consider the MIMO one ER scenario. The number of the antennas at the transmitter, IR and ER are  $N_t = 5$ ,  $N_i = 3$  and  $N_e = 3$ , resp.. The channels  $\tilde{\mathbf{H}}$  and  $\tilde{\mathbf{G}}$  are assumed to be quasi-static flat Rayleigh fading, and each element of them follows an independent complex Gaussian distribution with zero mean and a covariance specified by  $\tilde{\mathbf{L}}$ , where the transmitter is assumed to have the distances  $d_i = 9\text{m}$  and  $d_e = 7\text{m}$  away from IR and ER, resp... Also, we set other parameters as:  $\sigma_i^2 = \sigma_e^2 = -5\text{dBm}$ ,  $\eta = 0.8$ ,  $C_0 = 3\text{bit/s/Hz}$ ,  $\lambda = \mu = 15$ ,  $r^2 = 10$ ,  $\varepsilon_1 = \varepsilon_2 = \zeta = 10^{-3}$  and  $\mathbf{W}_{I0} = \mathbf{0}$ . Fig. 2 shows the convergence performance of Algorithm 1 when fixing  $P = 10\text{dB}$  and  $\alpha = 0.5$ . Two different initial feasible points after the feasibility check process of  $\mathbf{W}_{I0}$  are selected as the examples. The results show that the proposed EHM algorithm converges to the same harvested energy with different initializations (cf. Fig. 2(a)) and the initializations converge to optimal gradually after each iteration (cf. Fig. 2(b)). Fig. 2(c) and Fig. 2(d) show the convergence of the ellipsoid method adopted in updating the Lagrangian dual variables  $\lambda$  and  $\mu$ . During each update of variable  $\mathbf{W}_I$ , the ellipsoid method is introduced to obtain the relevant optimal dual variables  $\lambda^*$  and  $\mu^*$ . It is observed that the actual iteration numbers which fulfills the stop criterion is limited and less than the upper bound value (i.e.,  $8 \ln \left( \frac{r t_s}{\varepsilon} \right)$ ).

Then, we show the energy harvesting performance of our proposed algorithm and compare it with some existing methods and scenarios, namely, the iterative CVX (I-CVX) method, the GSVD method [9], the Plain-SVD method [16], the ER not serving as an eavesdropper (w/o eavesdropper) and the IR not possessing the ability to perform interference cancellation (w/o  $\mathbf{W}_E$  cancellation). Fig. 3 shows the average harvested energy of various methods and different scenario setups w.r.t. the transmitter power. It can be seen that our proposed algorithm performs identically to the I-CVX method and better than the other methods in terms of energy harvesting. Regarding the w/o eavesdropper case, since there is no intercept, less power is required to fulfill the rate constraint, which, thereby, resulting in more power transfer to ER. On the contrary, when the

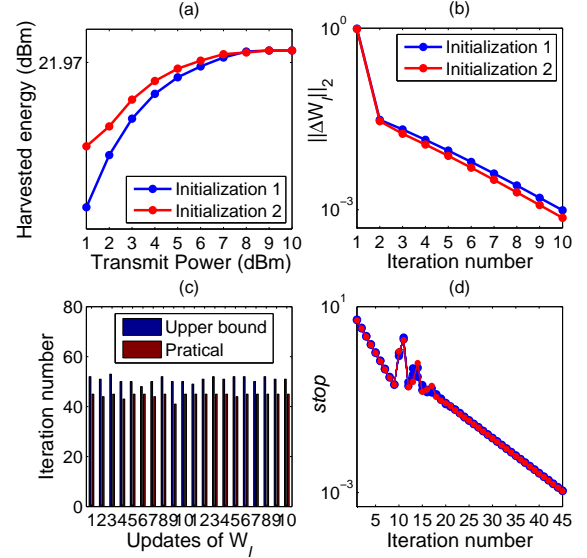


Fig. 2. Convergence analysis of Algorithm 1 for  $P = 10\text{dB}$  and  $\alpha = 0.5$ .

TABLE I  
AVERAGE RUNNING TIME (IN SECS.) VERSUS THE TRANSMIT POWER

Method	Transmit Power (dB)					
	3	6	9	12	15	18
Proposed	0.0875	0.1067	0.1339	0.1526	0.1877	0.2136
I-CVX	2.0628	2.4692	2.8700	3.3599	3.9975	5.8160

energy signals act as the disturbance to IR (i.e., the w/o  $\mathbf{W}_E$  cancellation case), more power is allocated to fulfill the secrecy capacity constraint, which, of course, leads to a worse energy harvesting performance for ER. Fig. 4 shows the harvested energy versus the secrecy capacity constraint with fixed  $P = 10\text{dB}$ . Again, our proposed algorithm yields the uniform performance with the I-CVX method and superior to other methods. In the w/o eavesdropper scenario, the harvested energy decreased gently with the increasing secrecy capacity requirement. The reason is that the lower bound of the desired rate of IR can be achieved easily so that a large proportion of transmitter power is assigned to transport energy for ER. In this figure, a trade-off design between the harvested energy at ER and secrecy capacity at IR is also provided for each method or scenario. Table I shows the average running time of our proposed and I-CVX methods. It reveals intuitively that our proposed algorithm owns a much lower computational complexity than I-CVX.

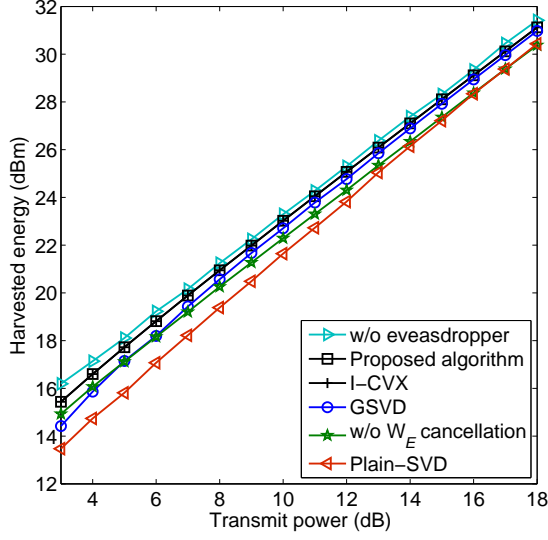


Fig. 3. Harvested energy versus the transmit power.

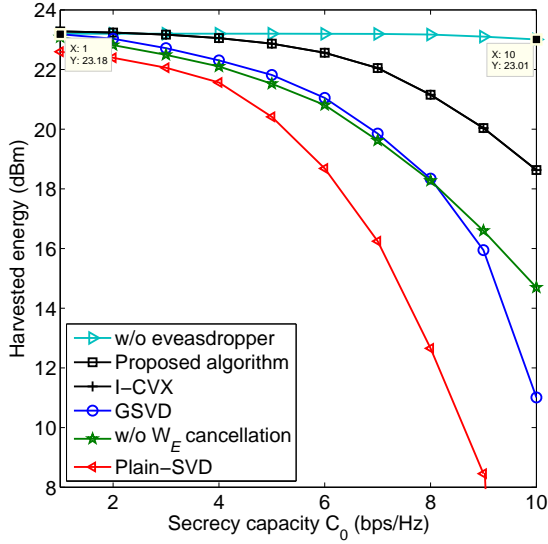


Fig. 4. Harvested energy versus the secrecy capacity.

### B. Multiple ERs WS-EHM

In this subsection, we consider the MIMO multiple ERs scenario. The channels  $\tilde{\mathbf{H}}$  and  $\tilde{\mathbf{G}}_k, k \in K$ , are the same as in previous subsection while  $d_i = 9m$  and  $d_{e,k} = 7m, \forall k$ . The other parameters are set as:  $N_t = 5$ ,  $N_i = 3$ ,  $N_{e,k} = 3, \forall k$ ,  $K = 3$ ,  $\sigma_i^2 = \sigma_{e,k}^2 = -5\text{dBm}$ ,  $\mu_k = 1, \forall k$ ,  $\eta_k = 0.8, \forall k$ ,  $\mu_t = 3$ ,  $q_1 = 0.1P$ ,  $\xi_1 = \xi_3 = 10^{-3}$ ,  $\xi_2 = 10^{-6}$  and  $\mathbf{W}_I^0 = \mathbf{W}_E^0 = \mathbf{V}^0 = (P/3N_t)\mathbf{I}$ . Fig. 5 shows the convergence performance of Algorithm 3 when fixing  $P = 10\text{dB}$  and  $R_0 = 3\text{bit/s/Hz}$  with any one channel realization. Fig. 5(a) and Fig. 5(b) illustrate the harvested energy versus the updates of  $t$  and iterations of variable  $(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V})$ , resp.. In Fig. 5(a), the curves from bottom to up represent the  $n$ th iteration of  $(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V})$  in turn. We can see that one circulation of

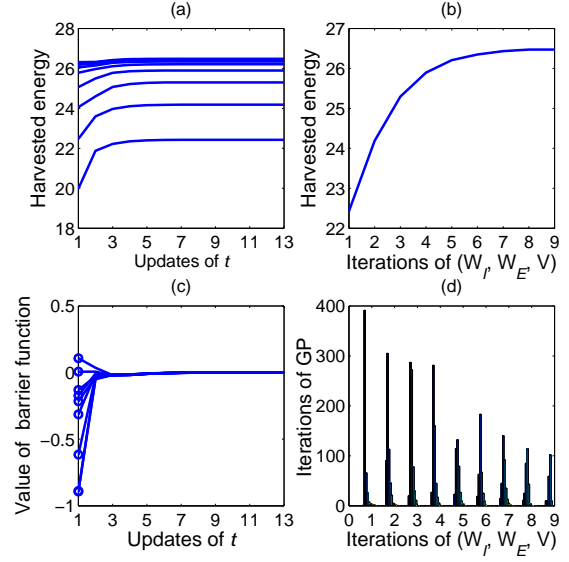


Fig. 5. Convergence analysis of Algorithm 3.

the updates of  $t$  is needed for every iteration of  $(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V})$ . Moreover, the harvested energy increases and converges to a fixed value finally. Also, to some degree, such a merit can be confirmed by Fig. 5(b). Similar to the curves in Fig. 5(a), different circles in Fig. 5(c) corresponding to different iteration of  $(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V})$ . From Fig. 5(c), we learn that  $x_k^n \geq 0$  is satisfied and the value of barrier function  $\frac{1}{t} \sum_{k=1}^K \ln(x_k^n)$  tends to zero with the update of  $t$ . Fig. 5(d) shows the innermost (GP method) iterations of Algorithm 3. Since the iterations of GP are limited and the resulting solutions derived from the GP method are in analytical forms, a fast and efficient operation of our proposed algorithm can be achieved.

After the illustration of the convergence of our proposed algorithm, we will focus on the investigation of these questions. Where are the advantages of our proposed algorithm? Why do we introduce AN in the multiple ERs case rather than the single ER case? With these two questions, we carried on the following simulation tests. In Fig. 6, we demonstrate the energy harvesting performance versus the transmit power when fixing  $R_0 = 3\text{bit/s/Hz}$ . To test the performance of our proposed algorithm (marked as Barrier-GP (proposed) in the legend), we compare it with the I-CVX method with the introduction of AN (marked as 'I-CVX w/  $\mathbf{V}$ '). From the plot, we can see that our proposed algorithm yields a consistent performance with the 'I-CVX w/  $\mathbf{V}$ ' method which is deemed to be a common approach in achieving the optimal objective value. However, from Table II, one can see that our proposed algorithm runs much faster than the 'I-CVX w/  $\mathbf{V}$ ' method. The speedup is within the range of 55 to 400 times. Then, to test the benefit of AN in promoting energy harvesting, different scenarios, such as multiple ERs and single ER, with and without AN, with and without  $\mathbf{W}_E$  cancellation at IR, etc., are taken into account in the plot. As an example, the legend 'multi-ER w/o  $\mathbf{V}$  w/  $\mathbf{W}_E$ ' represents the scenario of multiple ERs without the consideration of AN and the interference of

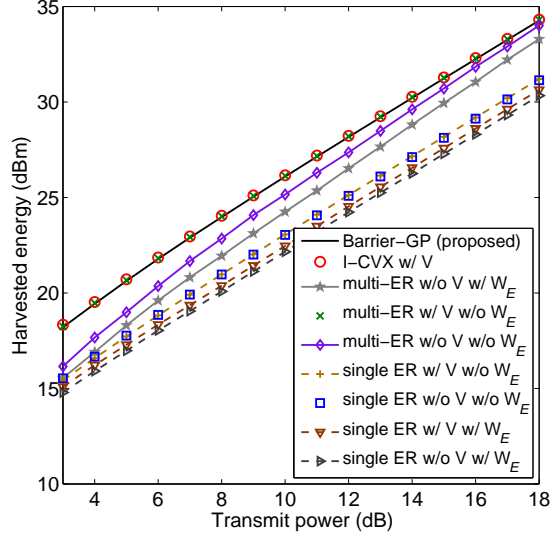


Fig. 6. Harvested energy versus the transmit power.

TABLE II  
AVERAGE RUNNING TIME (IN SECS.) VERSUS THE TRANSMIT POWER

Method	Transmit Power (dB)				
	3	6	9	12	15
Proposed	0.1036	0.1594	0.4741	0.9041	1.2387
I-CVX w/ V	41.8818	44.7600	49.7995	55.0403	68.3917

energy signals is not cancelled at IR. Through the comparison of ‘I-CVX w/ V’ and ‘multi-ER w/o V w/  $\mathbf{W}_E$ ’, we see a significant promotion in energy harvesting with the help of AN especially at the low power range. A similar result can also be affirmed from the comparison of ‘multi-ER w/ V w/o  $\mathbf{W}_E$ ’ and ‘multi-ER w/o V w/o  $\mathbf{W}_E$ ’. To the single ER scenarios, it witnesses an identical performance between ‘single ER w/ V w/o  $\mathbf{W}_E$ ’ and ‘single ER w/o V w/o  $\mathbf{W}_E$ ’. Though the ‘single ER w/ V w/  $\mathbf{W}_E$ ’ case shows a better performance than ‘single ER w/o V w/  $\mathbf{W}_E$ ’, the gain is almost negligible. To summarize, the above legends indicate that the AN plays a positive role in facilitating energy harvesting in the multiple ERs scenarios while not in the single ER scenarios.

Fig. 7 shows the harvested energy versus the secrecy rate when fixing  $\gamma$ . The methods and scenarios considered here are same to Fig. 6. From Fig. 7, we have the following three observations: First, the harvested energy decreases as the secrecy rate grows up. This is due to the fact that more transmit power is assigned to fulfill the higher secrecy rate requirement. Hence, the power be used for energy transfer become less. Second, for the multiple ERs scenarios, the energy-rate region enlarges apparently under the aid of AN especially in the ‘w/  $\mathbf{W}_E$ ’ case. Such an observation is consistent with our argument that the AN can both promote efficient energy transfer and guarantee secure communication in the multiple ERs MIMO system. Third, for the single ER scenarios, the energy-rate region changes slightly or even remain unchanged. This indicates the dispensability of AN in the single ER MIMO system.

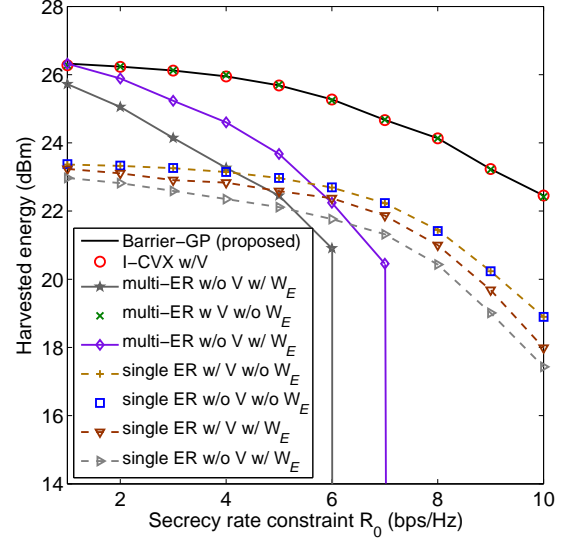


Fig. 7. Harvested energy versus the secrecy rate.

## VI. CONCLUSION

This paper has studied the energy harvesting problems in secure MIMO systems with one or multiple ERs. For each scenario, two types of IR, which is able or not able to remove the interference from energy signals, have been taken into account. In the scenario of single ER, we studied two EHM problems (P1 and P2) and proposed an iterative WF-SVD algorithm which results in the semi-closed form solutions for problem P1. In the scenario of multiple ERs, we studied the AN-aided WS-EHM problem and proposed a barrier-GP based block GS algorithm with proved KKT point convergence of this problem. Numerical results demonstrated the superior performance of our proposed algorithms. In the future, we will consider to extend our attention to more complex and practical systems, e.g., multiple multi-antenna IRs, the channel state information between the transmitter and receivers are imperfect, the transmitter is equipped with a large number of antennas, etc.

## APPENDIX A PROOF OF THEOREM 3.1

Simplify the expression of the Lagrangian in (14) with given  $\mathbf{W}_{I0}$ ,  $\lambda$  and  $\mu$ , we have

$$\mathcal{L}(\lambda, \mu, \mathbf{W}_I) = \lambda \ln(\mathbf{I} + \mathbf{H}^H \mathbf{W}_I \mathbf{H}) - \text{Tr}(\mathbf{Q} \mathbf{W}_I) + c_0, \quad (41)$$

where  $c_0 = \lambda \text{Tr}[(\mathbf{I} + \mathbf{G}^H \mathbf{W}_{I0} \mathbf{G})^{-1} \mathbf{G}^H \mathbf{W}_{I0} \mathbf{G}] + \mu P - \lambda \ln|\mathbf{I} + \mathbf{G}^H \mathbf{W}_{I0} \mathbf{G}| - \lambda R_0 \ln 2$ . Then, substitute  $\tilde{\mathbf{W}}_I$  for  $\mathbf{W}_I$  and let  $\tilde{\mathbf{W}}_I = \mathbf{Q}^{1/2} \mathbf{W}_I \mathbf{Q}^{1/2}$ , the equation (41) can be reformed as

$$\begin{aligned} \mathcal{L}(\lambda, \mu, \tilde{\mathbf{W}}_I) &= \lambda \ln(\mathbf{I} + \mathbf{H}^H \mathbf{Q}^{-1/2} \tilde{\mathbf{W}}_I \mathbf{Q}^{-1/2} \mathbf{H}) \\ &\quad - \text{Tr}(\tilde{\mathbf{W}}_I) + c_0. \end{aligned} \quad (42)$$

Do SVD about  $\mathbf{H}^H \mathbf{Q}^{-1/2}$ , i.e.,  $\mathbf{H}^H \mathbf{Q}^{-1/2} = \mathbf{U} \mathbf{\Lambda}^{1/2} \mathbf{V}^H$ , where  $\mathbf{U} \in \mathbb{C}^{N_i \times L}$  and  $\mathbf{V} \in \mathbb{C}^{N_t \times L}$ , and substitute it

into (42). Subsequently, the optimal analytic solution for the maximization of (42) can be achieved [32]:  $\tilde{\mathbf{W}}_I = \mathbf{V}\mathbf{D}\mathbf{V}^H$ , where  $\mathbf{D} = \text{diag}(p_1, p_2, \dots, p_L)$ . Now, let us seek for the solution of the diagonal element  $p_i, i = 1, \dots, L$ . Substituting  $\tilde{\mathbf{W}}_I = \mathbf{V}\mathbf{D}\mathbf{V}^H$  into (42), we have

$$\mathcal{L}(\lambda, \mu, \mathbf{D}) = \lambda \ln(\mathbf{I} + \mathbf{D}\mathbf{A}) - \text{Tr}(\mathbf{D}) + c_0. \quad (43)$$

Maximizing (43) about  $\mathbf{D}$ , combine with the standard WF power allocation method [32], the solution  $p_i$  can be obtained:  $p_i = \left(\lambda - \frac{1}{\delta_i^2}\right)$ ,  $i = 1, \dots, L$ . Thus,  $\tilde{\mathbf{W}}_I = \mathbf{Q}^{1/2}\mathbf{W}_I\mathbf{Q}^{1/2} = \mathbf{V}\mathbf{D}\mathbf{V}^H$  implies  $\mathbf{W}_I^* = \mathbf{Q}^{-1/2}\mathbf{V}\mathbf{D}\mathbf{V}^H\mathbf{Q}^{-1/2}$ , and the diagonal matrix  $\mathbf{D} = \text{diag}(p_1, \dots, p_L)$  is established with given  $p_i, i = 1, \dots, L$ .

Given  $\mathbf{W}_I$ , the dual variables  $\lambda$  and  $\mu$  can be optimized by minimizing the dual function (14b). Here, the subgradient-based approach, e.g., ellipsoid method [33], is employed to deal with the dual problem  $(\tilde{\mathbf{P}}\mathbf{I} - \mathbf{1} - \mathbf{D})$ . According to such a method, the subgradient, a key index during operations, of  $d(\lambda, \mu)$  at point  $[\lambda, \mu]$  can be defined as  $[\tilde{C}_{s1}^* - C_0, \alpha P - \text{Tr}(\mathbf{W}_I^*)]$ . Once the optimal dual solutions  $\lambda^*$  and  $\mu^*$  are figured out, subsequently, the optimal diagonal element  $p_i^* = \left(\lambda^* - \frac{1}{\delta_i^2}\right)$ ,  $i = 1, \dots, L$  will be derived as well as the optimal transmit information covariance  $\mathbf{W}_I^* = \mathbf{Q}^{-1/2}\mathbf{V}\mathbf{D}^*\mathbf{V}^H\mathbf{Q}^{-1/2}$  which converges to the optimal solution to subproblem  $(\tilde{\mathbf{P}}\mathbf{I} - \mathbf{1})$ . Thus, the proof of Theorem 3.1 is finished.

## APPENDIX B PROOF OF PROPOSITION 3.1

To have a better understanding of the update processes of the Lagrangian multipliers  $\lambda$  and  $\mu$  by applying the subgradient-based approach, we need the following lemma extracted from [33].

**Lemma B.1:** Considering the problem of maximizing a convex function  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  and an ellipsoid  $\mathbf{O}$  be described as

$$\mathbf{O} = \left\{ \mathbf{z} | (\mathbf{z} - \mathbf{x})^T \mathbf{P}^{-1} (\mathbf{z} - \mathbf{x}) \leq 1 \right\},$$

where  $\mathbf{P} \in \mathbf{S}_{++}^n$  gives the size and shape of  $\mathbf{O}$ ;  $\mathbf{x} \in \mathbb{R}^n$ , each element inside is nonnegative, is the center of the ellipsoid  $\mathbf{O}$ . Computing the subgradient of the objective function  $f$  at the  $j$ th iteration, it has  $\mathbf{s}^j = \partial f(\mathbf{x}^j)$ . The center point  $\mathbf{x}$  is updated as  $\mathbf{x}^j = \mathbf{x}^{j-1} - \frac{1}{n+1}\mathbf{P}^{j-1}\tilde{\mathbf{s}}$ , where  $\tilde{\mathbf{s}} = \frac{\mathbf{s}^{j-1}}{\sqrt{(\mathbf{s}^{j-1})^T \mathbf{P}^{j-1} \mathbf{s}^{j-1}}}$  and  $\mathbf{P}^j = \frac{n^2}{n^2-1} \left( \mathbf{P}^{j-1} - \frac{2}{n+1}\mathbf{P}^{j-1}\tilde{\mathbf{s}}\tilde{\mathbf{s}}^T \mathbf{P}^{j-1} \right)$ . Stopping the iteration once  $\sqrt{(\mathbf{s}^j)^T \mathbf{P}^j \mathbf{s}^j} \leq \varepsilon$ , where  $\varepsilon > 0$ , fulfills.

According to Lemma B.1, using the ellipsoid method, we assume that point  $\mathbf{x} = [\lambda, \mu]^T$  is the center of the ellipsoid  $\mathbf{O}$ . Then, given the initial matrix  $\mathbf{A}^0 \in \mathbf{S}_{++}^2$  which characterize the size and the orientation of the ellipsoid  $\mathbf{O}$ , the center point  $[\lambda, \mu]^T$  can be updated as  $[\lambda^j, \mu^j]^T = [\lambda^{j-1}, \mu^{j-1}]^T - \frac{1}{3}\mathbf{A}^{j-1}\tilde{\mathbf{s}}$ , where  $\tilde{\mathbf{s}} = \frac{\mathbf{s}}{\sqrt{\mathbf{s}^T \mathbf{A}^{j-1} \mathbf{s}}}$  and that  $\mathbf{s} = \left[ \frac{\partial \mathcal{L}(\mathbf{W}_I^*, \lambda^{j-1}, \mu^{j-1})}{\partial \lambda^{j-1}}, \frac{\partial \mathcal{L}(\mathbf{W}_I^*, \lambda^{j-1}, \mu^{j-1})}{\partial \mu^{j-1}} \right]^T =$

$\left[ \tilde{C}_{s1}^* - C_0, \alpha P - \text{Tr}(\mathbf{W}_I^*) \right]^T$ ,  $\mathbf{W}_I^*$  is the corresponding optimal value obtained from the  $j$ th iteration. The matrix  $\mathbf{A}$  can be updated as:  $\mathbf{A}^j = \frac{4}{3} \left( \mathbf{A}^{j-1} - \frac{2}{3}\mathbf{A}^{j-1}\tilde{\mathbf{s}}\tilde{\mathbf{s}}^T \mathbf{A}^{j-1} \right)$ . Moreover, the stopping criterion is thus formed as  $(\mathbf{s}^T \mathbf{A}^j \mathbf{s})^{1/2} \leq \varepsilon$ .

Next, let's discuss the convergence of the iteration. From the volume formula  $V(\mathbf{O}) = \beta_n \sqrt{\det(\mathbf{P})}$  in which  $\beta_n = \pi^{n/2} / \Gamma(n/2 + 1)$ , we elicit the volume whose size and shape are determined by matrix  $\mathbf{A} \in \mathbf{S}_{++}^2$ :  $V(\mathbf{O}) = \beta_2 \sqrt{\det(\mathbf{A})}$ . Suppose that our original ellipsoid  $\mathbf{O}^0$  contains the minimizer in its interior, i.e.,  $[\lambda^*, \mu^*]^T \in \mathbf{O}^0$ , the main idea of the ellipsoid method is to shrink the volume of the resulting ellipsoid of previous iteration such that  $\mathcal{L}_{best}(\mathbf{W}^*, \lambda^j, \mu^j) \leq \mathcal{L}(\mathbf{W}^*, \lambda^*, \mu^*) + \varepsilon$  satisfies, where  $[\lambda^j, \mu^j]^T$  is the center point of the ellipsoid  $\mathbf{O}^j$  at the  $j$ th iteration. Therefore, the convergence can be judged based on the change of the volume of ellipsoid  $\mathbf{O}^k$  which can be given as:

$$\begin{aligned} V(\mathbf{O}^j) &= \beta_2 \sqrt{\det(\mathbf{A}^j)} \\ &= \beta_2 \left[ \det \left( \frac{4}{3} \left( \mathbf{A}^{j-1} - \frac{2}{3}\mathbf{A}^{j-1}\tilde{\mathbf{s}}\tilde{\mathbf{s}}^T \mathbf{A}^{j-1} \right) \right) \right]^{\frac{1}{2}} \\ &= \beta_2 \left[ \det \left( \frac{4}{3} \left( \mathbf{A}^{j-1} - \frac{2}{3}\mathbf{A}^{j-1} \frac{\mathbf{s}\mathbf{s}^T}{\mathbf{s}^T \mathbf{A}^{j-1} \mathbf{s}} \mathbf{A}^{j-1} \right) \right) \right]^{\frac{1}{2}} \\ &\stackrel{(a)}{=} \frac{4}{3}\beta_2 [\det(\mathbf{A}^{j-1})]^{\frac{1}{2}} \left[ \det \left( \mathbf{I} - \frac{2}{3} \frac{\mathbf{s}\mathbf{s}^T}{\mathbf{s}^T \mathbf{A}^{j-1} \mathbf{s}} \mathbf{A}^{j-1} \right) \right]^{\frac{1}{2}} \\ &\stackrel{(b)}{=} \frac{4}{3}\beta_2 [\det(\mathbf{A}^{j-1})]^{\frac{1}{2}} \left[ 1 - \text{Tr} \left( \frac{2}{3} \frac{\mathbf{s}\mathbf{s}^T}{\mathbf{s}^T \mathbf{A}^{j-1} \mathbf{s}} \mathbf{A}^{j-1} \right) \right]^{\frac{1}{2}} \\ &= \frac{4}{3}\beta_2 [\det(\mathbf{A}^{j-1})]^{\frac{1}{2}} \left[ 1 - \frac{2}{3} \frac{\text{Tr}(\mathbf{s}\mathbf{s}^T \mathbf{A}^{j-1})}{\mathbf{s}^T \mathbf{A}^{j-1} \mathbf{s}} \right]^{\frac{1}{2}} \\ &= \frac{4}{3} \left[ 1 - \frac{2}{3} \right]^{\frac{1}{2}} \beta_2 [\det(\mathbf{A}^{j-1})]^{\frac{1}{2}} \\ &\stackrel{(c)}{<} e^{-1/4} V(\mathbf{O}^{j-1}) \\ &\stackrel{(d)}{<} e^{-J/4} V(\mathbf{O}^0), \end{aligned} \quad (44)$$

where (a) is due to the fact that  $\det(\mathbf{X}\mathbf{Y}) = \det(\mathbf{X})\det(\mathbf{Y})$ ; (b) is due to the fact that  $\det(\mathbf{I} - \mathbf{A}) = 1 - \text{Tr}(\mathbf{A})$  when  $\text{rank}(\mathbf{A}) \leq 1$ , here,  $\text{rank}(\mathbf{s}\mathbf{s}^T) = 1$  and  $\mathbf{A}^{j-1} \in \mathbf{S}_{++}^2$  lead to  $\text{rank}(\mathbf{s}\mathbf{s}^T \mathbf{A}^{j-1}) \leq 1$ ; (c) is result from a conclusion in [33] that  $\left(\frac{n}{n+1}\right)^{(n+1)/2} \left(\frac{n}{n-1}\right)^{(n-1)/2} < e^{-1/2n}$  and (d) is a recursive conclusion based on the derivations above it. From (44), we confirm that the volume of the ellipsoid  $\mathbf{O}^j$  decreases after each iteration. Therefore, the next work we need to do is to find the upper bound of our iterations. To begin with, let's define a ball

$$\mathbf{B} = \left\{ \mathbf{z} | \|\mathbf{z} - \mathbf{x}^*\| \leq \varepsilon/l_s \right\}, \quad (45)$$

where  $l_s = \max_{\mathbf{s} \in \partial \mathcal{L}(\mathbf{W}, \lambda, \mu)} \|\mathbf{s}\|$  is the maximum length of the subgradients over the initial ellipsoid  $\mathbf{O}^0$  and  $\mathbf{B} \subseteq \mathbf{O}^0$ . With this, we have  $\mathcal{L}(\mathbf{W}, \mathbf{z}) \leq \mathcal{L}(\mathbf{W}, \mathbf{x}^*) + \varepsilon$  when  $\mathbf{z} \in \mathbf{B}$ , while for  $\mathbf{z} \notin \mathbf{B}$ , it is deemed that  $\mathcal{L}(\mathbf{W}, \mathbf{z}) \geq \mathcal{L}(\mathbf{W}, \mathbf{x}^*) + \varepsilon$ . Since  $\mathbf{x}^*$  is defined as the final solution of the iteration, i.e., the achieved optimal solution, and ball  $\mathbf{B}$  is the infinitesimal region contains  $\mathbf{x}^*$ , we have  $\mathbf{B} \subseteq \mathbf{O}^j$  in iterations  $1, \dots, J$ . Hence,  $V(\mathbf{O}^j) \geq V(\mathbf{B})$ , which leads to

$$e^{-J/4} V(\mathbf{O}^0) \geq V(\mathbf{B}) = \beta_2 \left( \varepsilon/l_s \right)^2. \quad (46)$$

Taking the initial ellipsoid  $\mathbf{O}^0$  to be a ball, i.e.,  $\mathbf{x}^0 = [\lambda^0, \mu^0]^T = \mathbf{0}$  and  $\mathbf{A}^0 = r^2 \mathbf{I}$ . Then, do log operations of the first and last terms in (46), we have

$$-\frac{J}{4} + 2 \ln r \geq 2 \ln \frac{\varepsilon}{l_s}, \quad (47)$$

where  $V(\mathbf{O}^0) = \beta_2 r^2$ . Finally, the maximum iteration number can be expressed as

$$J \leq 8 \ln \frac{r l_s}{\varepsilon}. \quad (48)$$

So, we conclude that it takes no more than  $8 \ln \frac{r l_s}{\varepsilon}$  iterations to achieve  $\mathcal{L}(\mathbf{W}, \lambda^*, \mu^*)$  with error at most  $\varepsilon$ .

#### APPENDIX C PROOF OF PROPOSITION 4.2

Before the start of the proof, we need construct an optimization problem based on the subproblems  $(\bar{\text{P3}} - 1)$ ,  $(\bar{\text{P3}} - 2)$  and  $(\bar{\text{P3}} - 3A)$  solved in each iteration. Such a problem is formed as

$$(\bar{\text{P3}}) : \max_{\mathbf{W}_I, \mathbf{W}_E, \mathbf{V}} \text{EWS}(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V}) + \frac{1}{t} \sum_{k=1}^K \ln(x_k) \quad (49a)$$

$$\text{s.t.} \quad \text{Tr}(\mathbf{W}_I + \mathbf{W}_E + \mathbf{V}) \leq P, \quad (49b)$$

$$\mathbf{W}_I \succeq \mathbf{0}, \mathbf{W}_E \succeq \mathbf{0}, \mathbf{V} \succeq \mathbf{0}, \mathbf{T}_0 \succeq \mathbf{0}, \mathbf{T}_k \succeq \mathbf{0}, \forall k, \quad (49c)$$

where  $x_k = x_k(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V}, \{\mathbf{T}_k\}_{k=0}^K) = \max_{\mathbf{T}_0 \succeq \mathbf{0}, \mathbf{T}_k \succeq \mathbf{0}} \theta_I(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V}, \mathbf{T}_0) - \theta_{E,k}(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V}, \mathbf{T}_k) - R_0 \ln 2$ . Next, based on this, we complete the proof through two steps. First, we demonstrate that every limit point generate by the block GS method (cf. Algorithm 3) is a critical point of problem  $(\bar{\text{P3}})$ . Then, we show that the critical point of problem  $(\bar{\text{P3}})$  is also a critical point of the original problem (P3). Only then can our statement in Proposition 4.2 be ascertained.

In the first step, we first introduce the following lemma to support our preliminary convergence result.

**Lemma C.1:** ([28], Proposition 6) Consider such a problem:

$$\max_{\mathbf{x}} f(\mathbf{x}) \\ \text{s.t.} \quad \mathbf{x} \in \mathbf{X} = \mathbf{X}_1 \times \mathbf{X}_2 \times \cdots \times \mathbf{X}_m \subseteq \mathbb{R}^n, \quad (m > 2) \quad (50)$$

which can be solved iteratively by m-block GS method. Suppose that  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is pseudoconcave on  $\mathbf{X}$  and the level set  $\ell_{\mathbf{X}}^0 := \{\mathbf{x} \in \mathbf{X} : f(\mathbf{x}) \geq f(\mathbf{x}^0), \mathbf{x}^0 \in \mathbf{X}\}$  is compact. Then, it can be confirmed that the sequence  $\{\mathbf{x}^n\}$  produced by the m-block GS method has limit points and every limit point of  $\{\mathbf{x}^n\}$  is a critical point of (50).

It is easy to see that the objective (49a) is concave w.r.t. each variable, fixing the others. And, the level set  $\ell_{\mathbf{X}}^0$  is compact for every  $(\mathbf{W}_I^0, \mathbf{W}_E^0, \mathbf{V}^0, \mathbf{T}_0^0, \{\mathbf{T}_k^0\}_{k=1}^K) \in \mathfrak{N}$ , of which  $\mathfrak{N} = \{\mathbf{W}_I, \mathbf{W}_E, \mathbf{V}, \mathbf{T}_0, \{\mathbf{T}_k\}_{k=1}^K \mid x_k \geq 0, (49b), (49c)\}$ , on account of that the iterates  $(\mathbf{W}_I^n, \mathbf{W}_E^n, \mathbf{V}^n, \mathbf{T}_0^n, \{\mathbf{T}_k^n\}_{k=1}^K)$  are bounded (own to the total transmit power constraint) [28]. In consequence, invoking Lemma C.1, we conclude that  $(\mathbf{W}_I^n, \mathbf{W}_E^n, \mathbf{V}^n, \mathbf{T}_0^n, \{\mathbf{T}_k^n\}_{k=1}^K)$  have limit points and every

limit point produced by the block GS method is a critical point of problem  $(\bar{\text{P3}})$ .

In the second step, our endeavor is to prove that the critical point of problem  $(\bar{\text{P3}})$  is also a critical point of our original problem (P3). Suppose that is a critical point of problem (P3). The Lagrangian of problem  $(\bar{\text{P3}})$  can be written as

$$\begin{aligned} \mathcal{L}(\chi) = & \text{EWS}(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V}) + \frac{1}{t} \sum_{k=1}^K \ln(x_k) \\ & - v [\text{Tr}(\mathbf{W}_I + \mathbf{W}_E + \mathbf{V}) - P] + \text{Tr}(\Phi_{\mathbf{W}_I} \mathbf{W}_I) \\ & + \text{Tr}(\Phi_{\mathbf{W}_E} \mathbf{W}_E) + \text{Tr}(\Phi_{\mathbf{V}} \mathbf{V}) + \text{Tr}(\Phi_{\mathbf{T}_0} \mathbf{T}_0) \\ & + \sum_{k=1}^K \text{Tr}(\Phi_{\mathbf{T}_k} \mathbf{T}_k), \end{aligned}$$

where  $\chi$  denotes a collection of all the primal and dual variables of problem  $(\bar{\text{P3}})$ ,  $v \in \mathbb{R}_+$ ,  $\Phi_{\mathbf{W}_I} \in \mathbb{H}_+^{N_t}$ ,  $\Phi_{\mathbf{W}_E} \in \mathbb{H}_+^{N_t}$ ,  $\Phi_{\mathbf{V}} \in \mathbb{H}_+^{N_t}$ ,  $\Phi_{\mathbf{T}_0} \in \mathbb{H}_+^{N_i}$  and  $\Phi_{\mathbf{T}_k} \in \mathbb{H}_+^{N_e}$  are the dual variables associated with (49b),  $\mathbf{W}_I \succeq \mathbf{0}$ ,  $\mathbf{W}_E \succeq \mathbf{0}$ ,  $\mathbf{V} \succeq \mathbf{0}$ ,  $\mathbf{T}_0 \succeq \mathbf{0}$  and  $\mathbf{T}_k \succeq \mathbf{0}$ , resp.. Since  $(\mathbf{W}_I^*, \mathbf{W}_E^*, \mathbf{V}^*)$  maximizes  $\mathcal{L}(\chi)$  for given  $\mathbf{T}_0^*$  and  $\{\mathbf{T}_k^*\}_{k=1}^K$ , its gradients w.r.t.  $\mathbf{W}_I$ ,  $\mathbf{W}_E$  and  $\mathbf{V}$  vanishes at  $(\mathbf{W}_I^*, \mathbf{W}_E^*, \mathbf{V}^*)$ , see (51) where  $x_k^* = x_k(\mathbf{W}_I^*, \mathbf{W}_E^*, \mathbf{V}^*, \{\mathbf{T}_k^*\}_{k=0}^K)$ ,  $\Omega^* = \mathbf{H}[\mathbf{I} + \mathbf{H}^H(\mathbf{W}_I^* + \mathbf{W}_E^* + \mathbf{V}^*)\mathbf{H}]^{-1}\mathbf{H}^H$  and  $\Psi^* = \mathbf{G}_k[\mathbf{I} + \mathbf{G}_k^H \mathbf{V}^* \mathbf{G}_k]^{-1}\mathbf{G}_k^H$ . Defining

$$\lambda_k^*(t) = \frac{1}{t x_k^*}, \forall k$$

we claim that  $\lambda_k^*(t)$  is the dual variable w.r.t. the  $k$ th secrecy rate constraint in problem (P3) and know that  $\lambda_k^*(t) > 0$  due to  $x_k^* > 0, \forall k$ . Moreover, as  $t \rightarrow \infty$ , we have  $\lambda_k^*(t) x_k^* = 0, \forall k$ . Besides, according to Lemma 4.1, the final results of the introduced auxiliary variables  $\mathbf{T}_0$  and  $\{\mathbf{T}_k\}_{k=1}^K$  take the forms, resp.:

$$\begin{aligned} \mathbf{T}_0^* &= [\mathbf{I} + \mathbf{H}^H(\mathbf{W}_E^* + \mathbf{V}^*)\mathbf{H}]^{-1}, \\ \mathbf{T}_k^* &= [\mathbf{I} + \mathbf{G}_k^H(\mathbf{W}_I^* + \mathbf{V}^*)\mathbf{G}_k]^{-1}, \forall k. \end{aligned} \quad (52)$$

Let  $\phi_k(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V}) = C_I(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V}) - C_{E,k}(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V}) - R_0 \ln 2, \forall k$ , substituting (52) into  $x_k(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V}, \{\mathbf{T}_k\}_{k=0}^K)$ , we gain  $\phi_k(\mathbf{W}_I^*, \mathbf{W}_E^*, \mathbf{V}^*) = x_k(\mathbf{W}_I^*, \mathbf{W}_E^*, \mathbf{V}^*, [\mathbf{I} + \mathbf{H}^H(\mathbf{W}_E^* + \mathbf{V}^*)\mathbf{H}]^{-1}, \{[\mathbf{I} + \mathbf{G}_k^H(\mathbf{W}_I^* + \mathbf{V}^*)\mathbf{G}_k]^{-1}\}_{k=1}^K) = x_k^*$ . Thus, (51) can be rewritten as

$$\nabla_{\mathbf{W}_I} \text{E}_{\text{WS}}^* + \sum_{k=1}^K \lambda_k^*(t) \nabla_{\mathbf{W}_I} \phi_k^* - v^* \mathbf{I} + \Phi_{\mathbf{W}_I}^* = \mathbf{0}, \quad (53a)$$

$$\nabla_{\mathbf{W}_E} \text{E}_{\text{WS}}^* + \sum_{k=1}^K \lambda_k^*(t) \nabla_{\mathbf{W}_E} \phi_k^* - v^* \mathbf{I} + \Phi_{\mathbf{W}_E}^* = \mathbf{0}, \quad (53b)$$

$$\nabla_{\mathbf{V}} \text{E}_{\text{WS}}^* + \sum_{k=1}^K \lambda_k^*(t) \nabla_{\mathbf{V}} \phi_k^* - v^* \mathbf{I} + \Phi_{\mathbf{V}}^* = \mathbf{0}, \quad (53c)$$

$$\lambda_k^*(t) \phi_k^* = 0, \forall k, \quad (53d)$$

where  $\text{E}_{\text{WS}}^* = \text{EWS}(\mathbf{W}_I^*, \mathbf{W}_E^*, \mathbf{V}^*)$ ,  $\phi_k^* = \phi_k(\mathbf{W}_I^*, \mathbf{W}_E^*, \mathbf{V}^*)$  and (53d) is result from the previous conclusion  $\lambda_k^*(t) x_k^* = 0, \forall k$ . Since  $(\mathbf{W}_I^*, \mathbf{W}_E^*, \mathbf{V}^*, \{\mathbf{T}_k^*\}_{k=0}^K)$  is the optimal solution of problem  $(\bar{\text{P3}})$ , it must satisfy the KKT conditions of

$$\left\{ \begin{array}{l} \nabla_{\mathbf{W}_I} \mathcal{L}(\chi) = \sum_{k=1}^K \sigma_{e,k}^2 \mu_k \eta_k \mathbf{G}_k \mathbf{G}_k^H + \frac{1}{t} \sum_{k=1}^K \frac{1}{x_k^*} \underbrace{(\Omega^* - \mathbf{G}_k \mathbf{T}_k^* \mathbf{G}_k^H)}_{\nabla_{\mathbf{W}_I} x_k^*} - v^* \mathbf{I} + \Phi_{\mathbf{W}_I}^* = \mathbf{0}, \\ \nabla_{\mathbf{W}_E} \mathcal{L}(\chi) = \sum_{k=1}^K \sigma_{e,k}^2 \mu_k \eta_k \mathbf{G}_k \mathbf{G}_k^H + \frac{1}{t} \sum_{k=1}^K \frac{1}{x_k^*} \underbrace{(-\mathbf{H} \mathbf{T}_0^* \mathbf{H}^H + \Omega^*)}_{\nabla_{\mathbf{W}_E} x_k^*} - v^* \mathbf{I} + \Phi_{\mathbf{W}_E}^* = \mathbf{0}, \\ \nabla_{\mathbf{V}} \mathcal{L}(\chi) = \sum_{k=1}^K \sigma_{e,k}^2 \mu_k \eta_k \mathbf{G}_k \mathbf{G}_k^H + \frac{1}{t} \sum_{k=1}^K \frac{1}{x_k^*} \underbrace{(-\mathbf{H} \mathbf{T}_0^* \mathbf{H}^H + \Omega^* + \Psi^* - \mathbf{G}_k \mathbf{T}_k^* \mathbf{G}_k^H)}_{\nabla_{\mathbf{V}} x_k^*} - v^* \mathbf{I} + \Phi_{\mathbf{V}}^* = \mathbf{0}, \end{array} \right. \quad (51)$$

problem (P3), i.e., as listed in (54).

$$\left\{ \begin{array}{l} (51), \nabla_{\mathbf{T}_0} \mathcal{L}(\chi) = \mathbf{0}, \nabla_{\mathbf{T}_k} \mathcal{L}(\chi) = \mathbf{0}, \forall k, \\ v^* [\text{Tr}(\mathbf{W}_I^* + \mathbf{W}_E^* + \mathbf{V}^*) - P] = 0, v^* \geq 0, \\ \text{Tr}(\Phi_{\mathbf{W}_I}^* \mathbf{W}_I^*) = 0, \text{Tr}(\Phi_{\mathbf{W}_E}^* \mathbf{W}_E^*) = 0, \text{Tr}(\Phi_{\mathbf{V}}^* \mathbf{V}^*) = 0, \\ \text{Tr}(\Phi_{\mathbf{T}_0}^* \mathbf{T}_0^*) = 0, \text{Tr}(\Phi_{\mathbf{T}_k}^* \mathbf{T}_k^*) = 0, \forall k, \\ \Phi_{\mathbf{W}_I}^* \succeq \mathbf{0}, \Phi_{\mathbf{W}_E}^* \succeq \mathbf{0}, \Phi_{\mathbf{V}}^* \succeq \mathbf{0}, \\ \Phi_{\mathbf{T}_0}^* \succeq \mathbf{0}, \Phi_{\mathbf{T}_k}^* \succeq \mathbf{0}, \forall k. \end{array} \right. \quad (54)$$

Finally, by combining (53) and parts of the KKT conditions in (54), we note that  $(\mathbf{W}_I^*, \mathbf{W}_E^*, \mathbf{V}^*)$  must satisfy the following KKT conditions:

$$\left\{ \begin{array}{l} \nabla_{\mathbf{W}_I} E_{\text{WS}}^* + \sum_{k=1}^K \lambda_k^*(t) \nabla_{\mathbf{W}_I} \phi_k^* - v^* \mathbf{I} + \Phi_{\mathbf{W}_I}^* = \mathbf{0}, \\ \nabla_{\mathbf{W}_E} E_{\text{WS}}^* + \sum_{k=1}^K \lambda_k^*(t) \nabla_{\mathbf{W}_E} \phi_k^* - v^* \mathbf{I} + \Phi_{\mathbf{W}_E}^* = \mathbf{0}, \\ \nabla_{\mathbf{V}} E_{\text{WS}}^* + \sum_{k=1}^K \lambda_k^*(t) \nabla_{\mathbf{V}} \phi_k^* - v^* \mathbf{I} + \Phi_{\mathbf{V}}^* = \mathbf{0}, \\ \lambda_k^*(t) \phi_k^* = 0, \lambda_k^*(t) \geq 0, \forall k, \\ v^* [\text{Tr}(\mathbf{W}_I^* + \mathbf{W}_E^* + \mathbf{V}^*) - P] = 0, v^* \geq 0, \\ \text{Tr}(\Phi_{\mathbf{W}_I}^* \mathbf{W}_I^*) = 0, \text{Tr}(\Phi_{\mathbf{W}_E}^* \mathbf{W}_E^*) = 0, \text{Tr}(\Phi_{\mathbf{V}}^* \mathbf{V}^*) = 0, \\ \Phi_{\mathbf{W}_I}^* \succeq \mathbf{0}, \Phi_{\mathbf{W}_E}^* \succeq \mathbf{0}, \Phi_{\mathbf{V}}^* \succeq \mathbf{0}, \\ (\mathbf{W}_I^*, \mathbf{W}_E^*, \mathbf{V}^*) \in \mathbf{Z}, \end{array} \right. \quad (55)$$

where  $\mathbf{Z} = \{(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V}) | \phi_k(\mathbf{W}_I, \mathbf{W}_E, \mathbf{V}) \geq 0, \text{Tr}(\mathbf{W}_I + \mathbf{W}_E + \mathbf{V}) \leq P, \mathbf{W}_I \succeq \mathbf{0}, \mathbf{W}_E \succeq \mathbf{0}, \mathbf{V} \succeq \mathbf{0}\}$ . Clearly, the conditions in (55) are exactly the KKT conditions of problem (P3).

## REFERENCES

- [1] N. Tesla, "Apparatus for Transmitting Electrical Energy," *Patent US* 1,119, 732, Dec. 1914.
- [2] R. Want, "'Enabling ubiquitous sensing with RFID,'" *IEEE Computer*, vol. 37, pp. 84-86, Apr. 2004.
- [3] L. R. Varshney, "Transporting information and energy simultaneously," in *Proc. 2008 IEEE Int. Symp. Inf. Theory*, pp. 1612-1616.
- [4] P. Grover and A. Sahai, "Shannon meets Tesla: wireless information and power transfer," in *Proc. 2010 IEEE Int. Symp. Inf. Theory*, vol. 37, pp. 84-86, Apr. 2004.
- [5] D. W. K. Ng, R. Schober, "Max-min fair wireless energy transfer for secure multiuser communication systems," *Information Theory Workshop (ITW), 2014 IEEE*, pp. 326-330, Nov. 2014.
- [6] D. W. K. Ng, E. S. Lo, and R. Schober, "Robust Beamforming for Secure Communication in Systems with Wireless Information and Power Transfer," *IEEE Trans. on Wireless Comm.*, vol. 13, no. 8, pp. 4599-4615, August 2014.
- [7] W. Wu and B. Wang, "Robust Secrecy Beamforming for Wireless Information and Power Transfer in Multiuser MISO Communication System," *EURASIP Journal on Wireless Communications and Networking*, 2015:161, June 2015.
- [8] R. Zhang and C. Ho, "MIMO broadcasting for simultaneous wireless information and power transfer," in *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 1989-2001, May 2013.
- [9] W. Wu and B. Wang, "Efficient Transmission Solutions for MIMO Wiretap Channels With SWIPT," in *Communications Letters, IEEE*, vol. 13, no. 8, pp. 4599-4615, August 2014.
- [10] A. D. Wyner, "The wire-tap channel," *The Bell System Technical J.*, vol. 54, pp. 1355-1387, Oct. 1975.
- [11] I. Csiszar and J. Korner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 339-348, May 1978.
- [12] S. K. Leung-Yan-Cheong and M. E. Hellman, "The Gaussian wiretap channel," *IEEE Trans. Inf. Theory*, vol. 24, no. 4, pp. 451-456, July 1978.
- [13] S. Loyka and C. D. Charalambous, "On optimal signaling over secure MIMO channels," in *IEEE International Symposium on Information Theory*, pp. 443-447, July 2012.
- [14] S. Loyka and C. D. Charalambous, "An Algorithm for Global Maximization of Secrecy Rates in Gaussian MIMO Wiretap Channels," available at <http://arxiv.org/abs/1504.03725>. April, 2015.
- [15] Q. Li and W.-K. Ma, "Spatially selective artificial-noise aided transmit optimization for MISO multi-Eves secrecy rate maximization," *IEEE Trans. Signal Process.*, vol. 61, no. 10, pp. 2704-2717, May 2013.
- [16] Q. Li et al., "Transmit solutions for MIMO wiretap channels using alternating optimization," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1714-1727, Sep. 2013.
- [17] Hui-Ming Wang, Feng Liu and Xiang-Gen Xia, "Joint Source-Relay Precoding and Power Allocation for Secure Amplify-and-Forward MIMO Relay Networks," *IEEE Trans. on Information Forensics and Security*, vol. 9, no. 8, pp. 1240-1250, Aug. 2014.
- [18] X. Zhou, R. Zhang, and C. K. Ho, "Wireless Information and Power Transfer: Architecture Design and Rate-Energy Tradeoff," in *IEEE Trans. Commun.*, vol. 61, no. 11, pp. 4754-4767, Nov. 2013.
- [19] J. Xu, L. Liu, and R. Zhang, "Multiuser MISO Beamforming for Simultaneous Wireless Information and Power Transfer," *IEEE Intern. Conf. on Acoust. Speech and Signal Process.*, pp. 4754-4758, May 2013.
- [20] L. Liu, R. Zhang, and K.-C. Hua, "Secrecy wireless information and power transfer with MISO beamforming," *IEEE Trans. Signal Process.*, vol. 62, no. 7, pp. 1850-1863, Apr. 2014.
- [21] Muhammad R. A. Khandaker and Kai-Kit Wong, "Max-min energy based robust secure beamforming for SWIPT," *Communications (ICC), 2015 IEEE International Conference on*, pp. 166-171, June 2015.
- [22] Qingjiang Shi, Weiqiang Xu, Jinsong Wu, Enbin Song and Yaming Wang, "Secure Beamforming for MIMO Broadcasting With Wireless Information and Power Transfer," in *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2841-2853, May 2015.
- [23] Haiyang Zhang, Chunguo Li, Yongming Huang and Luxi Yang, "Secure Beamforming for SWIPT in Multiuser MISO Broadcast Channel With Confidential Messages," in *Communications Letters, IEEE*, vol. 19, no. 8, pp.1347-1350, Aug. 2015.
- [24] Meng Zhang and Yuan Liu, "Energy Harvesting for Physical-Layer Security in OFDMA Networks," in *Information Forensics and Security, IEEE Transactions on*, vol. 11, no. 1, pp. 154-162, Jan. 2016.
- [25] R. Feng, Q. Li and Q. Zhang, "Robust Secure Transmission in MISO Simultaneous Wireless Information and Power Transfer System," *IEEE Trans. Veh. Tech.*, vol. 64, no. 1, pp. 400-405, Jan. 2015.
- [26] J. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optim. Methods Softw.*, vol. 11, pp. 625-653, 1999, (webpage and software) <http://sedumi.ie.lehigh.edu/>
- [27] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming," Apr. 2011, <http://cvxr.com/cvx>
- [28] L. Grippio and M. Scianrone, "On the convergence of the block

- nonlinear Gauss-Seidel method under convex constraints,” *Operation research letter*, vol. 26, pp. 127-136, 2000.
- [29] D. Bertsekas, *Nonlinear Programming*, Belmont, MA: Athena Scientific, 1999.
  - [30] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge, UK: Cambridge University Press, 2004.
  - [31] Y. Liang, G. Kramer, H. V. Poor, and S. Shamaï (Shitz), “Compound wire-tap channels,” in *Proc. 45th Annu. Allerton Conf. Commun., Control, Comput.*, Sep. 2007, pp. 136-143.
  - [32] T. Cover and J. Thomas, *Elements of Information Theory*. Wiley, 1991.
  - [33] S. Boyd, “Convex optimization II,” Stanford University. Available: <http://stanford.edu/class/ee364b/lectures.html>
  - [34] B. R. Marks and G. P. Wright, “A general inner approximation algorithm for non-convex mathematical programs,” *Operations Research*, vol. 26, no. 4, pp. 681-683, 1978.
  - [35] D. P. Palomar, “Convex primal decomposition for multicarrier linear MIMO transceivers,” *IEEE Trans. Signal Process.*, vol. 53, no. 12, pp. 4661-4674, Dec. 2005.